PART B

Case study & Applications

4 Case Study

"Of all the experiences that you live, consider only the wisdom that they contain"
M. Twain

In this second part of the thesis surveying jobs, programming, testing and simulating activities are shown They have been developed in order to test models and techniques that have been described in PART A: Theory & Fundamentals.

4.1 LOCATION

In order to test grouping parameters, their validation field and applicability, as well as models and theories it has been necessary to introduce a real case study on which testing and programming has been developed. Since it was a "controlling field", it has been chosen an existing road where the design project was known and, thanks to this previous knowledge, it has been also possible to check any possible error in all performed elaboration.
Figure 4.1 shows the highway segment that was selected as the case study. The segment is included between two roundabouts (E & W) located along the National Route n.23 that links Turin to Sestriere (Italy), two of the main competition sites for the 2006 Winter Olympic Games. Prior to the event, the Piedmont Region financed the construction of this road segment, which was designed according to current Italian Geometric Policy for new roads (Ministero delle Infrastrutture e dei Trasporti, 2001). The specifications impose the use of spirals (clothoid) for the transition between tangents and circular curves.
The analysis has been performed using a cartographic coordinate system referred to the UTM32-WGS84 reference system. In this area, the combined value of mapping linear deformation and height scale factor is about 0.99970: the differences between distances represented in this mapping system and the corresponding actual distances are about 30 cm/km (19 in./mi). These values are negligible in comparison with the instrumental accuracy of low cost sensors (more or less 1 m – 3.2 ft), hence the UTM coordinate system can, in practice, be intended as an isometric coordinate system (mapping distance = real distance) for the purpose of this work.
The travelled way is composed of two lanes with a total width of 7.5 m (24.6 ft), and two shoulders each 1.5 m (4.9 ft) in width. The geometric characteristics of the twenty-five elements considered are listed in Table 4.1. For each, the table provides the length and the station from the first point of the element 1. Furthermore, the table gives the geometric characteristics of all the design alignment element of the road:

- length, starting station, radius and center coordinates for circular arcs,
- length, starting station and angular coefficient for tangents, and finally
- Length, starting station and scale factors for clothoids.

#	ELEMENT	ANG. COEFF	LENGTH m(ft)	STARTING STATION m(ft)	RADIUS m(ft)	CENTER COORDINATES (UTM WGS84) Xc	Yc	A m(ft)
1	CIRCULAR CURVE		49.28 (161.68)	0.00 (0.00)	80 (262.47)	3662.9589	1987.6140	
2	TRANSITION (CLOTHOID)		37.81 (124.06)	49.28 (161.68)				55 (180.45)
3	TANGENT	1.824	43.71 (143.41)	87.09 (285.74)				
4	TRANSITION (CLOTHOID)		46.55 (152.71)	130.80 (429.14)				160 (524.93)
5	CIRCULAR CURVE		36.71 (120.44)	177.35 (581.85)	550 (1804.46)	3068.4871	2215.6895	
6	TRANSITION (CLOTHOID)		46.55 (152.71)	214.06 (702.29)				160 (524.93)
7	TANGENT	1.309	531.67 (1744.32)	260.60 (855.00)				
8	TRANSITION (CLOTHOID)		46.55 (152.71)	792.27 (2599.31)				160 (524.93)
9	CIRCULAR CURVE		76.63 (251.41)	838.82 (2752.02)	550 (1804.46)	3591.7729	1088.2028	
10	TRANSITION (CLOTHOID)		46.55 (152.71)	915.45 (3003.43)				160 (524.93)
11	TANGENT	2.189	75.23 (246.82)	961.99 (3156.14)				
12	TRANSITION (CLOTHOID)		46.55 (152.71)	1037.22 (3402.96)				160 (524.93)
13	CIRCULAR CURVE		90.67 (297.47)	1083.77 (3555.66)	550 (1804.46)	2540.3165	1434.6003	
14	TRANSITION (CLOTHOID)		46.55 (152.71)	1174.44 (3853.14)				160 (524.93)
15	TANGENT	1.243	607.43 (1992.87)	1220.99 (4005.84)				
16	TRANSITION (CLOTHOID)		62.63 (204.16)	1828.42 (5998.72)				185 (606.95)
17	CIRCULAR CURVE		504.48 (1655.11)	1890.64 (6202.88)	550 (1804.46)	2125.4018	919.0062	
18	TRANSITION (CLOTHOID)		62.23 (204.16)	2395.12 (7857.99)				185 (606.95)
19	TANGENT	-0.138	2.23 (7.32)	2457.35 (8062.15)				
20	TRANSITION (CLOTHOID)		62.23 (204.16)	2459.58 (8069.46)				185 (606.95)
21	CIRCULAR CURVE		105.59 (346.42)	2521.81 (8273.62)	550 (1804.46)	1911.2389	-162.4647	
22	TRANSITION (CLOTHOID)		62.23 (204.16)	2627.40 (8620.04)				185 (606.95)
23	TANGENT	0.174	733.08 (2405.11)	2689.62 (8824.20)				
24	TRANSITION (CLOTHOID)		57.80 (189.63)	3422.70 (11229.31)				170 (557.74)
25	CIRCULAR CURVE		111.66 (366.34)	3480.50 (11418.94)	500 (1640.42)	1121.3365	-249.3504	
				3592.16 (11785.27)				

Table 4.1 - Case study alignment elements table

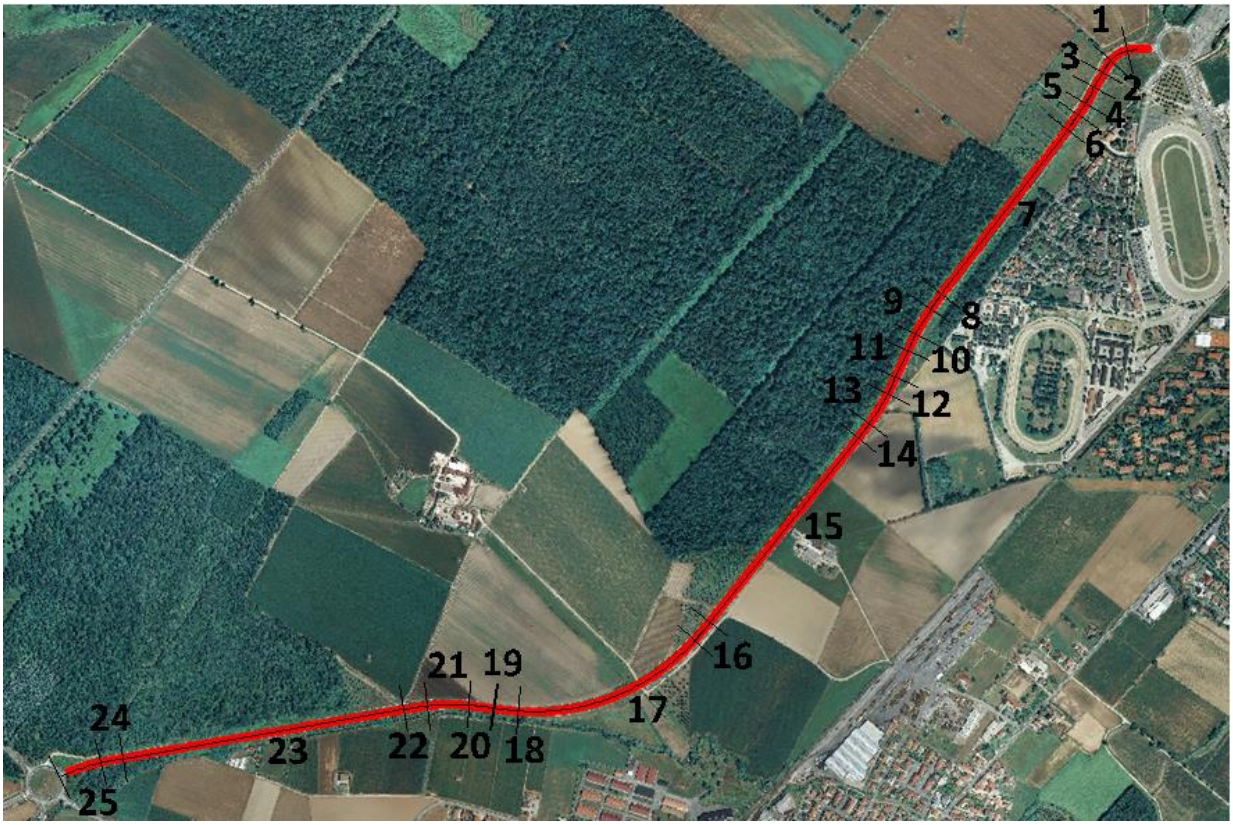


Figure 4.1 – Case study aerial image: highway segment along the Italian National Route n.23

4.2 Mobile Mapping survey

With the MMS concept as a starting point, in 2007 the Geomatics Research Group of the Politecnico di Torino designed and built a universal MMS. This system consisted of a metallic bar that hold up to four GNSS antennas, three inertial sensors (Crossbow IMU 700CA, Crossbow IMU 400CC, and XSens MTi), three web-cameras (Logitech Quickcam Pro 9000, with a resolution of up to 2 megapixels), and a terrestrial laser scanner (Sick LMS100).

With the aim of developing a new MMS to be used specifically in the field of alignment detection, Road Engineering and Geomatics research groups worked together on the development of a new apparatus (Figure 4.2). It is composed of one integrated platform GPS/IMU (XSens MTi-G, Table 4.2) and a webcam (Logitech Quickcam Pro 9000).

The new MMS prototype presented in Figure 4.2 has been fitted out with low cost sensors with a medium level of accuracy. Three other geodetic GNSS receivers Leica GX1230+GNSS were mounted on the same bar (Figure 4.2). These last mentioned sensors are most likely redundant but they were used to verify and assess the quality of the new system. This MMS required some "homemade" modifications prior to use. In particular, a dedicated data acquisition software program and a calibration system were developed to obtain geo-referenced and calibrated data.

Specification - IMU	
Angular Rate	
Range Roll, Pitch, Yaw (°/s)	± 300
Bias: Roll, Pitch (°/s)	± 0.5
Bias: Yaw (°/s)	± 1.0
Resolution (°)	0.05
Acceleration	
Range X/Y/Z, m/s^2 (ft/sec^2)	± 50 (± 164.04)
Bias: X/Y/Z, m/s^2 (ft/sec^2)	± 0.02 (± 0.06)
Resolution, m/s^2 (ft/sec^2)	0.0098 (± 0.0322)
Update Rate, Hz	512
Specification - GPS	
Raw Measurement	L1 Frequency, C/A code
# Channels	50
Max Update Rate, Hz	4
Operating Temperature, °C (°F)	from -40 to 85 (from -40 to 185)

Table 4.2 - X-Sens MTi-G Spefications

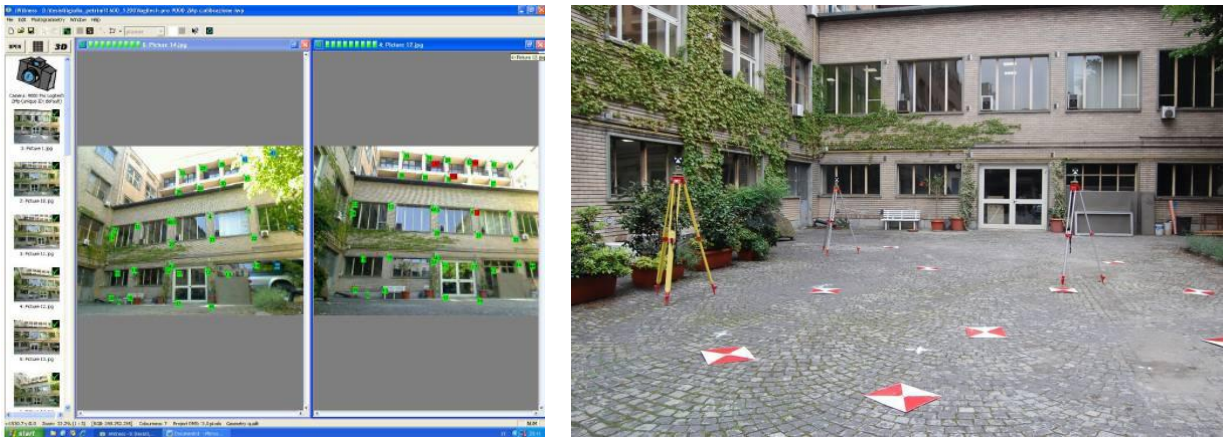
4.2.1 System calibration

The calibration of optical lenses, the definition of reference frame transformation, the synchronization of time devices and the integration between sensors have been considered and estimated in order to derive data useful for geometrical analysis.
As well known, every optical lens suffers from radial and tangential distortions that must be eliminated. In this case, a dedicated calibration field was realized, with several photos taken from different vantage points and angulations. The acquired images have been processed with iWitness, which is a software program dedicated to the estimation of lens distortions using the auto-calibration algorithm through a selection of homologous points (Figure 4.3).

Figure 4.2 - The new, low cost prototype MMS with sensors.
(1) GNSS receivers Leica GX1230+ (2) GPS/IMU XSens MTi-G,
(3) Logitech Quickcam Pro 9000 webcam, (4) LMS-110 Sick laser scanner

4.2.2 Data Acquisition

Each sensor on the MMS has a local reference system. The relationships (three shifts, three rotations and one scale factor) between each local system and the mapping reference system (East, North and Elevation) are the fundamental elements required to assign, in the final stage, the absolute coordinates to each pixel of the image. Using a calibration field and a total station, the positions of several targets were estimated in each local reference system with respect to the absolute reference system (UTM projection). Geometric relationships were estimated by means of a homographic transformation (Ewald, 1971).



(a) (b)

Figure 4.3 - Calibration process
(a) Screenshot of the iWitness software used for lens calibration (b) Calibration field

Time synchronization is a key aspect of every MMS because it permits the coordination of all the collected data. It is necessary to have one common time because each sensor has its own internal clock, with its own drift, accuracy, and stability. Any discrepancies would result in data (images, positions, attitudes) not being synchronized. The GPS time was taken as a reference. For this purpose the software GEOWASP (Lingua, Balbo, Giorgi, & Piras, 2009) was used. The Geomatics Research Group in Politecnico di Torino and the non-profit association Information Technology for Humanitarian Assistance, Cooperation, and Action (ITHACA) developed the software in the C #language and dedicated Windows API routines were also included.

The package is composed of two modules. The first one is dedicated to the synchronization between PC-time and GPS-time. At the same time, GPS time and PC time are written to a text file. It is also possible to include the instantaneous coordinate and the time of recording of a particular point of interest. The second GEOWASP module acquires the images and allows the management of up to six webcams in simultaneous operation. The webcam configuration (zoom, resolution, light conditions, etc.) could be changed using the webcam's own software (Logitech® Webcam Software). In this case, the resolutions used were equal to 960x720 pixels, with the number of frames per second equal to 8-10.

The IMU data were collected using a sample rate of 50 Hz through the MT Manager software developed by XSens (Figure 4.4). The data are stored in a proprietary format (.mtb) and can be easily converted into an ASCII file using a dedicated tool. The MTi-G sensor directly gives a loosely coupled solution (GPS/IMU solution), which is composed of three-dimensional positions (latitude, longitude and ellipsoidal height), accelerations and velocities along the three main axes (X,Y,Z) and the related accuracies. These solutions are stored in a text file but, unfortunately, the raw data (IMU and GPS) are not stored since the coupled GPS/IMU data is not available in post-processing.

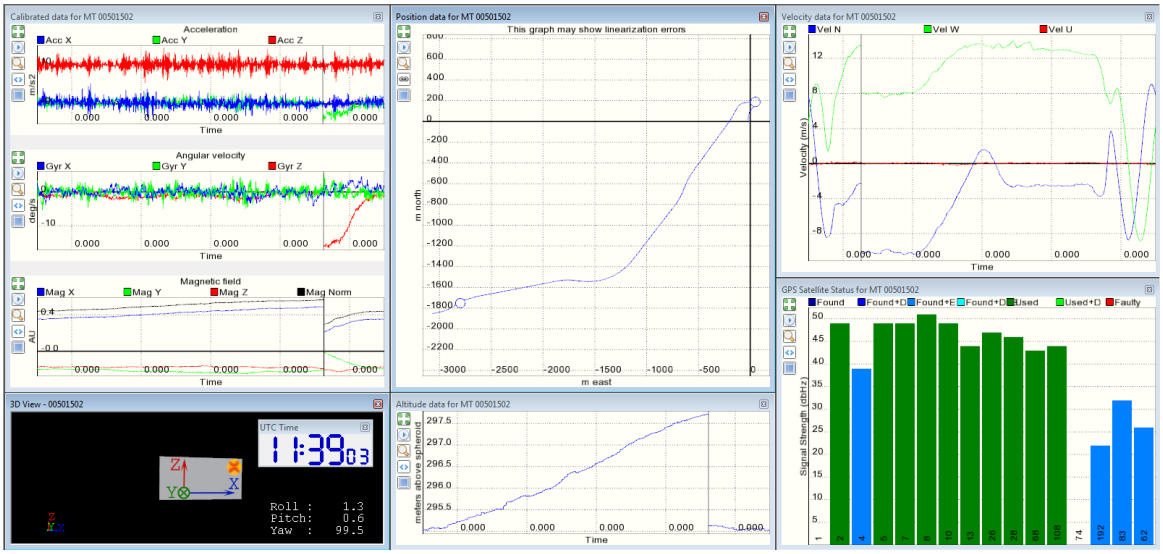


Figure 4.4 - Screen panel of MT Manager of the IMU-GPS sensor

4.2.3 Data validation and post-processing strategies

Before starting with the image analysis and the related lines extraction, an IMU data and image validation procedures were carried out.
First, the GPS/IMU stand-alone trajectory was compared with the GNSS one, which was obtained with a post-processing kinematic (PPK) solution. This comparison was realized differentiating the three-dimensional positions, epoch-by-epoch. Accuracies and precisions (considering a confidence level equal to 1 standard deviation) of this comparison are for horizontal components 1.10 m and 0.42 m (3.60 ± 1.38 ft) respectively, while for vertical ones are 1.18 m and 0.75 m (3.87 ± 2.46 ft). A geodetic receiver with a rate equal to 1 Hz acquired the unique GNSS raw data and they were used to estimate the stern position of the system with a level of accuracy to within centimeters. The GNSS post-processing was performed using the commercial package Leica Geo Office v5 (LGO), considering the GNSS reference station installed at the Politecnico di Torino as master station.
After that, an image quality control check was carried out to evaluate the brightness, contrast, and focus of each image. Excluding some singular cases, the majority of the images can be used for our purposes, even in cases where the focus is not particularly stable but the lines are well detectable.

Data post-processing involved the implementation of the following strategies:

- attitudes and positions of the vehicle have been obtained directly by GPS/IMU integrated sensor; in this case the navigation solution has been re-sampled to 10 Hz and the position of each image has been defined by means of an interpolation; and
- Attitudes of the vehicle have been estimated by GPS/IMU integrated sensor and position is derived from GNSS processing (PPK solution); in this case the attitudes have been re-sampled to 10 Hz and the position and attitude of each image has been defined by means of an interpolation.

4.2.4 Image analysis

Before proceeding with the image data processing, a pre-analysis of the images was carried out, with the aim of improving the quality of the images by means of the following two steps. Firstly the grayscale conversion and secondly the variation of the contrast level based on the equalization histogram.

The extraction of the pavement markings has been realized using this procedure. The RGB image has been converted into an intensity image (gray levels=16 bit/pixel) which has been stretched to emphasize white lines, borders and other radiometric edges. The Canny Filter edge extractor (2004) has been applied and the binary image obtained has been concealed using a mask image to force the feature recognition only in the road area. The Hough transform (2004) has been used to extract some features (i.e. lines) which have been automatically merged and classified in several layers (right/ left borders of lateral/white centerline, lane boundaries, etc.). Then, a robust algorithm based on LMS has been used to decimate the extracted data generating a more correct and precise set of data. This procedure has been completely implemented in Matlab language, where some internal toolboxes devoted to image analysis have been considered (i.e Hough function, edge function). In this context, no mathematical formulation will be described because the used algorithms are well documented in the literature (2004).

The steps of this procedure are shown in Figure 4.6. After that, the geo-referencing of the images, using the attitudes and positions available, have been performed with the plotting results saved to a dxf/dwg file.

From the image analysis the vehicle trajectory, the centerline, the roadside, the lateral line have been extracted; each single line has been detected in both road direction, so that in order to derive the road central axis it has been possible to perform an average between both directions:

- for each single point of the going track has been performed a research in the return track for all points inside a space window between 7 and 40 meters. This specific window research value has been selected keeping with reference the road width and possible mistakes and vehicle misleading;
- referring to all points inside this first selection, distances have been evaluated and the shortest one has led to the closest point on the opposite direction;
- repeating this operation with all points of the going track an average track has been obtained;
- Repeating this operation with all the lanes detected in the image analysis (centerlines, roadside, lateral lines, and trajectories) several "central average lines" has been obtained, each one originated from different road elements [Figure 4.5].

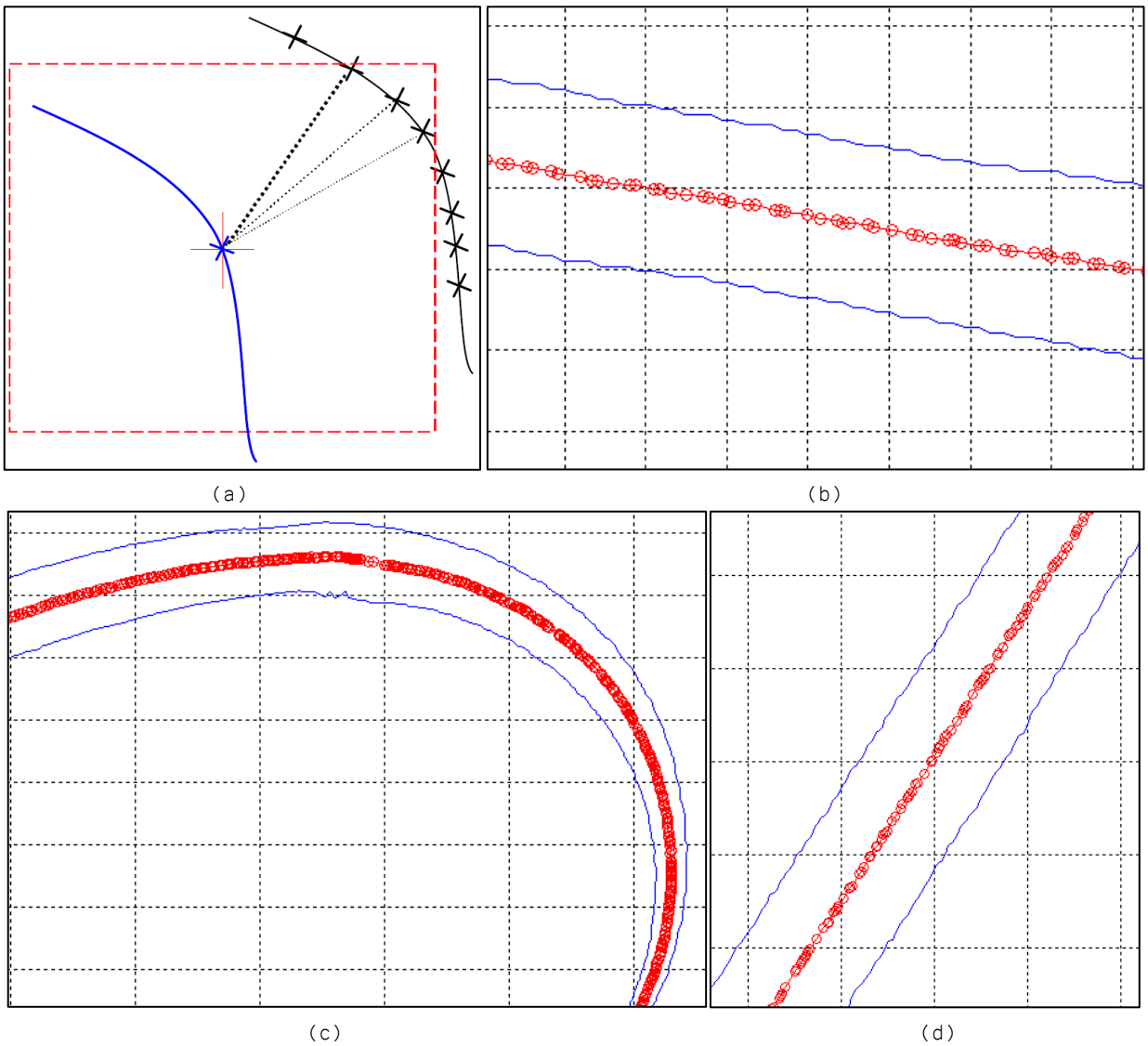


Figure 4.5 - Average trajectory process:
(a) Closest point research (b, d) Tangent section average extraction (c) Circular curve average extraction

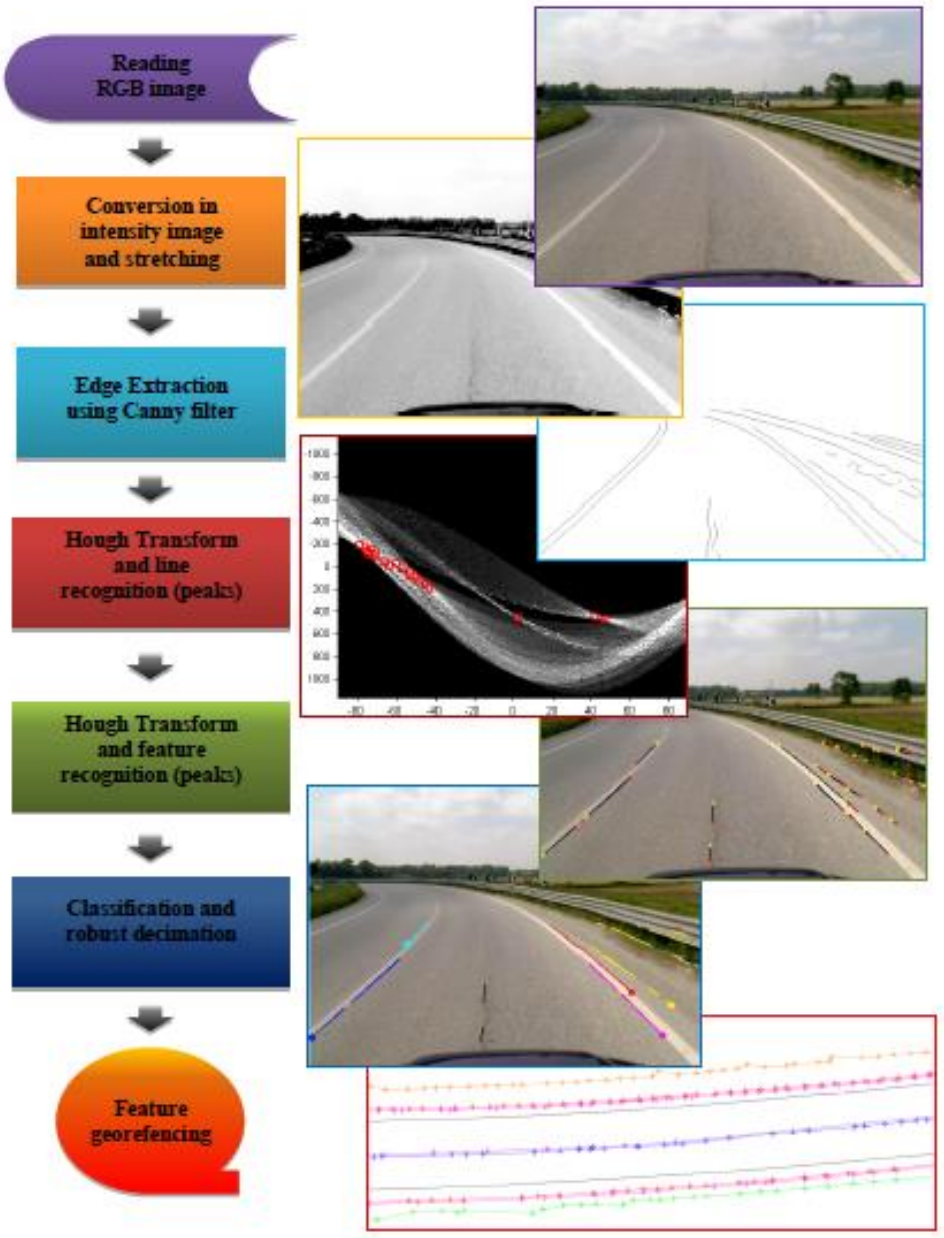


Figure 4.6 – Flow chart of road entities recognition procedure

4.3 AERIAL GIS PHOTOS

To evaluate the possibility of using aerial images to get alignments information, a visual and manual sampling process has been applied to oriented images of the interested area. Using GIS software (in this specific case, ArcGIS®) the centerline of the case study road has been manually extracted, with a different spacing related to each specific geometrical element of the alignment. In particular, 5 m spacing has been adopted while approaching, going through and exiting from curves, while 10 m spacing has been used along tangents.

After this manual operation, a shapefile has been produced and the list of nodes of the polyline generated has been extracted from the software database and imported into MatLab® for further elaboration, like any other experimental track.

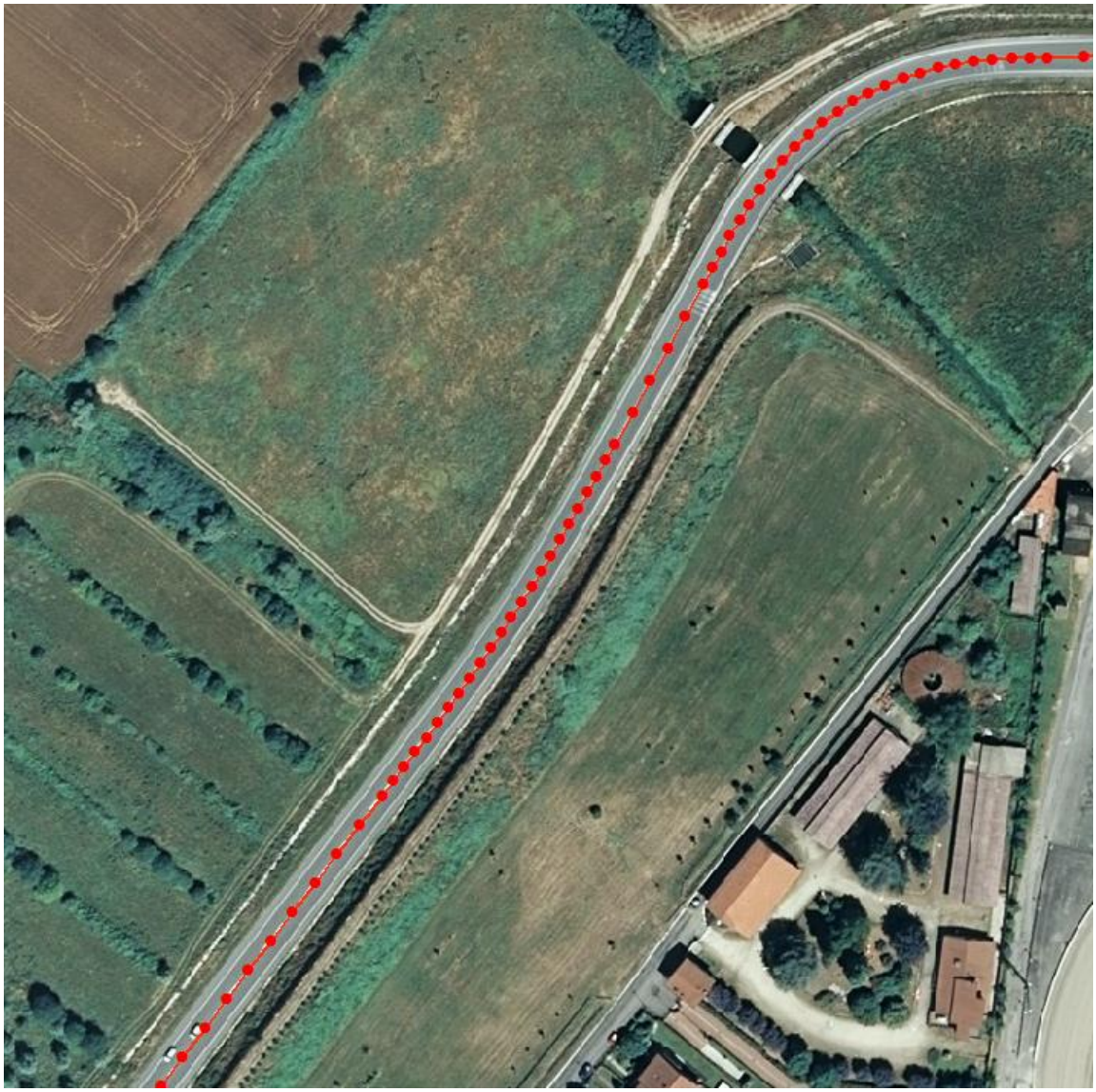


Figure 4.7 – GIS manual sampling on aerial image

4.4 DESIGN PROJECT DISCRETIZATION

As already discussed in [2.4], in this work an innovative approach methodology that deals with regression analyses from a different perspective was adopted. In all previous works generally information about road alignments starting from a specific database are obtained through surveying method. In fact, this is exactly what has been done up to now also in this work: it has been decided to investigate on fitting and regression models and techniques and a case study has been chosen in order to test this techniques.

Nevertheless, testing algorithms and technologies on real survey databases can lead to possible mistakes during quality assessments: as a matter of fact not only the quality of the algorithm/model is responsible for the results fitness, but also the data performances are fundamental. So that it can happen that the quality assessment of an algorithm is also affected by inaccuracies on data source.

To avoid this problem, it has been decided to propose a new perspective, **creating a new database completely theoretical**, obtained from the design project of the case study. That is the only reason why in this research report only that specific case study has been selected: the selected road was at the same time able to provide in-field mobile mapping surveys, aerial pictures and design project information, giving the possibility to work from different point of view.

Starting with this idea, it has been decided to transform the design project into a discrete point series. This discretization operation has been made with the Autocad© command *Draw – Point – Measurement.* This specific command gives the possibility to draw a series of points on a specified element keeping a predefined distance between points. The procedure transforms the CAD composition of the design file into a single geometrical entity (from separated tangents, clothoids, and circular curves into a single polyline), than the series of points has been generated. It has been chosen to set spacing equal to 0.5 m.
In a second step, all points' coordinates have been extracted from the dwg file into an external txt file.

After these data extractions from different origins, it is available a starting database composed like the one summarized in Table 4.3:

SOURCE ROAD ELEMENT	TRACK NAME	# PTS	POSITIONING REFERENCE
Centerline manual Sampling	GIS_CntLine	349	GIS Aerial Image
Round trip vehicle trajectory average	IMU_Traj	2129	Inertial Measurement Unit (IMU)
Round trip roadside average	IMU_GigLine	2132	Inertial Measurement Unit (IMU)
Round trip centerline average	IMU_CntLine	1338	Inertial Measurement Unit (IMU)
Round trip vehicle trajectory average	GPS_Traj	240	Global Positioning System (GPS)
Round trip lateral line average	GPS_LatLine	1847	Global Positioning System (GPS)
Round trip roadside average	GPS_CigLine	2002	Global Positioning System (GPS)
Round trip centerline average	GPS_CntLine	1397	Global Positioning System (GPS)
Design-based exact alignment	PROJ	7019	Design Project

Table 4.3 – Starting tracks details

All following elaborations have been performed starting from this initial database, and this process is shown in the next chapter.

c5 Database & coding

"Measuring the evolution of a programming process from the code lines number is like measuring the evolution in the construction of aircrafts by their weight"
B.Gates

All programming and coding processes have been performed on MatLab and it has been decided to save the entire database as a structure tree (mat file), called **TRACKS**. So, basically, all works has been applied on the database TRACKS with external algorithms especially developed to work in it: these algorithms have been named with a progressive number to remember the progress from one algorithm to the other.
Summarizing, all algorithms have been called with a name like "**work_NN**" (where NN is the progressive number), while all the results have been stored in the variable **TRACKS**, composed by a structure tree.

5.1 STARTING PROCEDURES

First, the database has been the homogenized on all tracks regarding the start point and the end one. Actually, since all different tracks were originated from different sources and with different techniques, they have different starting and final points, giving problems in terms of comparison between length and stations. That is the reason why it has been operated a truncation to all tracks in order to have starting and final points closest as possible to each other.
To save these starting and final points two variables have been built: **ENDPTS** and **STRPTS**, where have been stored the starting and final points (coordinates and index position) for each original track of the database.

In addition to that, all tracks have been stored splitting in separate vectors **East**, **North** and **Elevation**, and two more information have been added: **corresponding stations** to each point and the **total length** of the track.
Another important operation that has been made on the tracks is the sampling with different spacing of the design-based tracks.
Since the discretization of design project has been performed with spacing equal to 0.5 m, in order to simulate different tracks origins (in terms of surveying techniques and instruments) it has been decided to generate some new tracks with different spacing. To get these new tracks, the original 0.5 m track has been re-sampled with several new spacings: **1 m**, **2 m**, **5 m**, **10 m**, **15 m**, and **20 m**.
After these starting procedures the database is composed like the scheme in Figure 5.1.

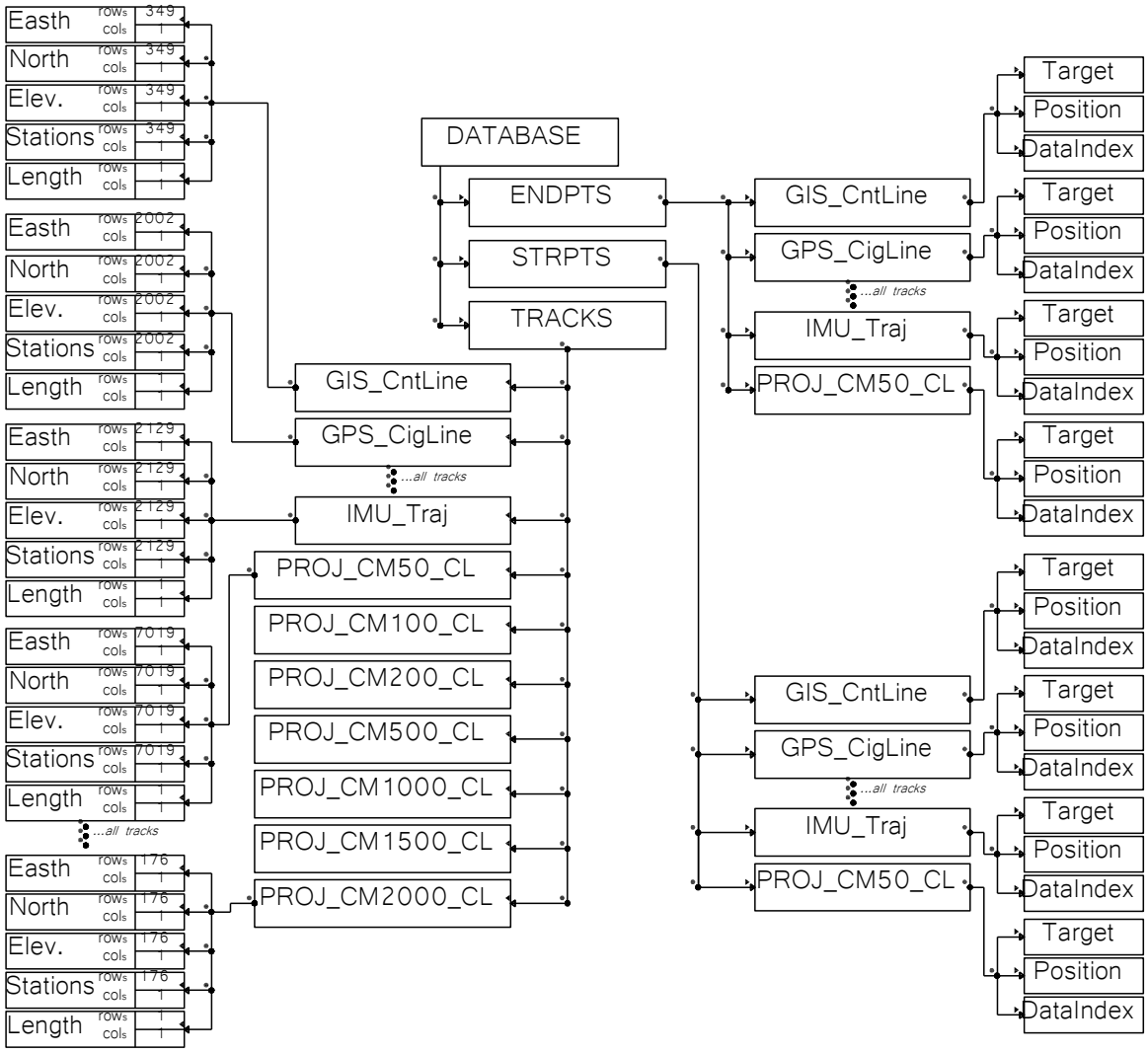


Figure 5.1- Database structure

5.2 CODING AND DATABASE CONSTRUCTION

5.2.1 Random error generation

As already said in the previous paragraph, the project-based track has been re-sampled with different spacing to simulate various surveying techniques and instruments. After this process, different levels of accuracy naturally linked to instruments and techniques have been adopted. In order to keep the innovative perspective that has been illustrated in [4.4], **different levels of accuracies for each design track have been simulated** to recreate a range of possible accuracies that can be obtained with several surveying instruments.
With direct reference to in-field track, a **random error** with different maximum magnitude has been assumed: 2 cm, 5 cm, 10 cm, 20 cm, and 40 cm.

The error generation procedure has been applied with the MatLab© function **randn** in: after specifying dimensions, it returns an n-by-n matrix of pseudorandom normal values. Thanks to this function, inside each single track a different level of "inaccuracy" was added, creating a new series of additional tracks. Each single point of the sequence was moved slightly from its original position by applying two noise terms (Δx and Δy) through the implementation of a sub-routine in the algorithm. 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

mean equal to zero and a standard deviation equal to 1 (i.e., Gaussian distribution). The errors were then scaled to a specified accuracy (or maximum error). Figure 5.2 and Figure 5.3 contain the logical flow-chart for this operation, coupled with the database update situation. As it's visible in the picture, each single track obtained from the project (varying the spacing) has generated up to other 6 tracks where the maximum error is different from a minimum to a maximum value (from 2 cm to 40 cm). These specific boundaries have been selected referring to the accuracy magnitude expected to be obtained from in-field surveys, where the bigger imprecision is equal to the definition of aerial pictures in GIS environment, where each single pixel cover a 40 cm wide real area. Random errors generated through this procedure are represented graphically in Figure 5.4, while the detailed view of each specific case is available in attach. "*Random errors detailed view*"
The corresponding MatLab© code is available in attach. "*Random error generation*".

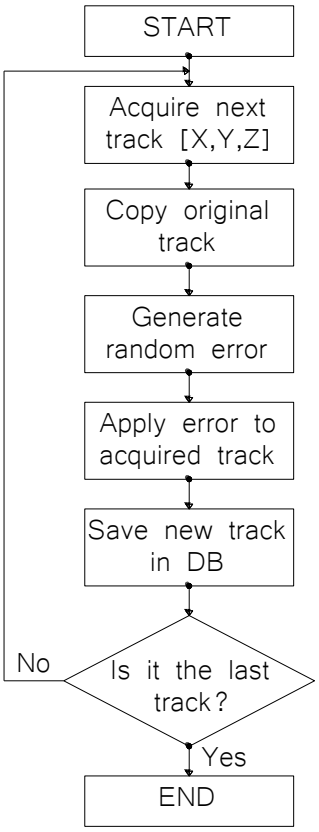


Figure 5.2 - Error random generation process

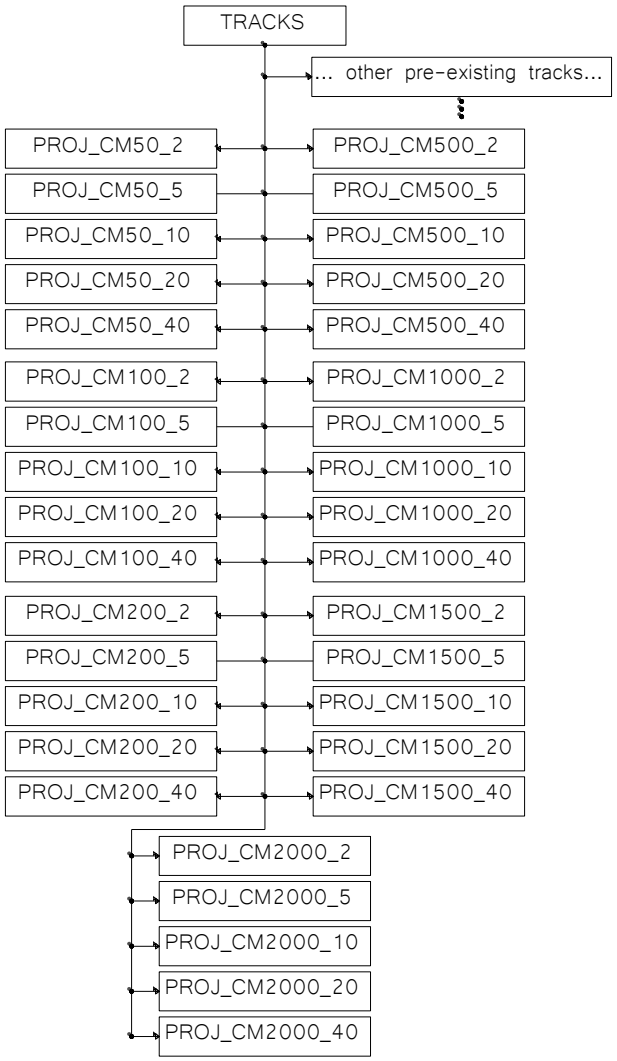


Figure 5.3 - Updated database

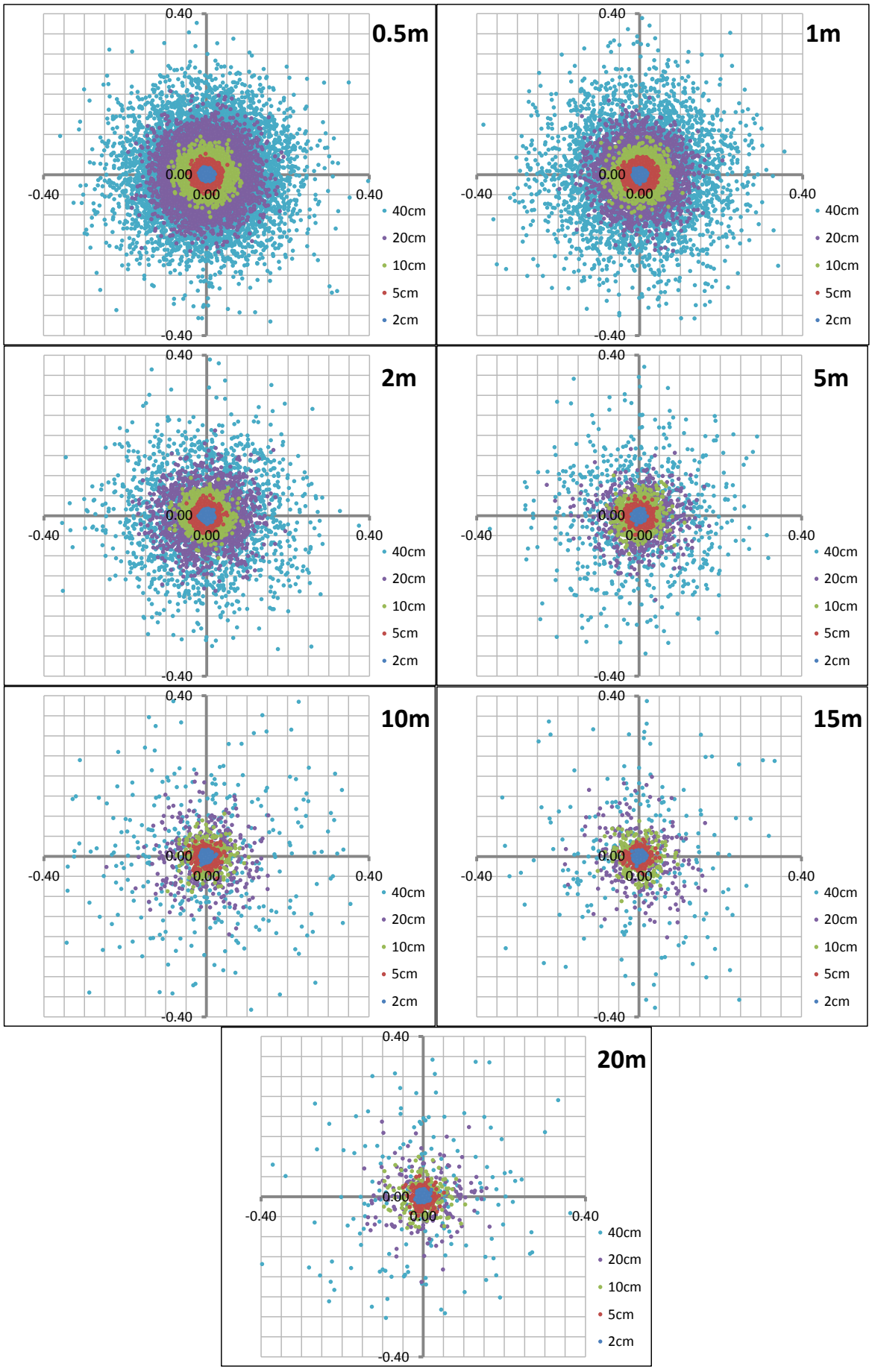


Figure 5.4 – Random error distribution around design discretization points

5.2.2 Deviation angle evaluation

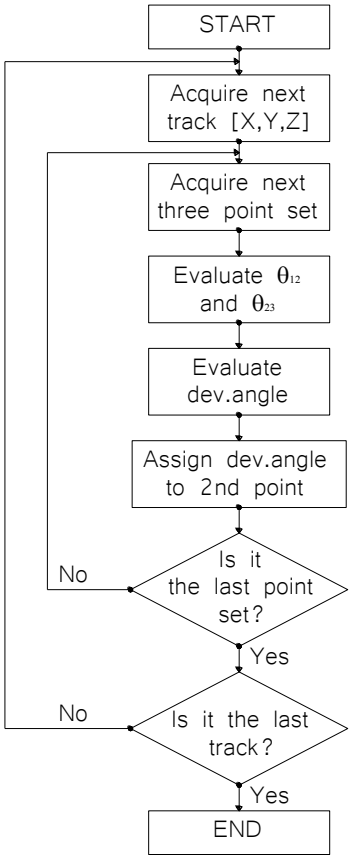


Figure 5.6 – Deviation angle evaluation

As discussed in [2.2] one of the most used parameters to perform the "grouping operation" to a road survey is the local deviation angle. Also in this research work it has been evaluated, and a specific algorithm has been built in MatLab language to calculate and store this parameter. It has been chosen to evaluate local deviation angle using azimuth methodology [see 2.2.2]. The logical process of the algorithm is shown in Figure 5.6, while the algorithm itself is available in the attachment: "*Local deviation angle*". The deviation angle is evaluated for each single point of each track (excluded the first and the last point, where a "NaN" – Not a Number – value is stored), and are stored directly in each track, adding a new variable called **Dev.Angle.**

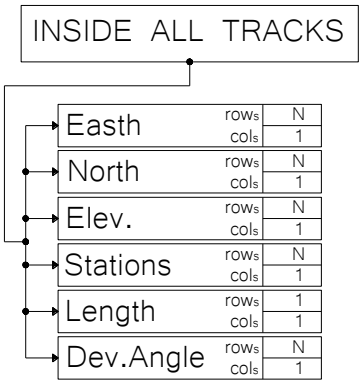


Figure 5.5 – Updated database detail with deviation angle evaluation

5.2.3 Local curvature with polynomial fitting

Polynomial technique [2.3.3] has been adopted to evaluate local curvature for all tracks, characterized with different spacing and accuracies. In order to control polynomial function it has been limited its degree to 2nd and 3rd. Various fitting mobile windows spacing have been adopted (they represent the number of points which are locally taken into account to fit the polynomial function). It is intended that every fitted polynomial is "assigned" to the central point of the mobile fitting window, and the evaluated curvature will be assigned as well to the central point.
Mobile fitting window spacing are the following:

- 3 points (2nd degree only),
- 5 points (2nd and 3rd degree),
- 9 points (2nd and 3rd degree),
- 13 points (2nd and 3rd degree),
- 17 points (2nd and 3rd degree), and
- 21 points (2nd and 3rd degree).

Each track has been characterized with 11 curvature combinations, evaluated in each single point (edge points have been excluded due to the impossibility to establish a mobile fitting window). Eleven vectors have been created, stored inside each single track. The logical process of the algorithm is shown in Figure 5.7, while the algorithm itself is available in attach. "*Local curvature with polynomial fitting*".

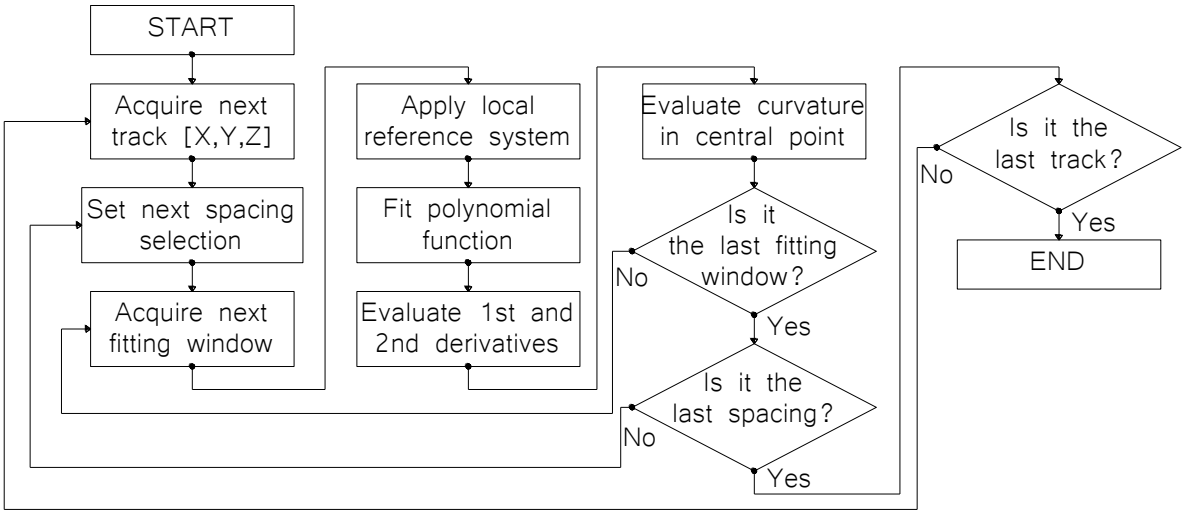


Figure 5.7 – Local curvature polynomial fitting

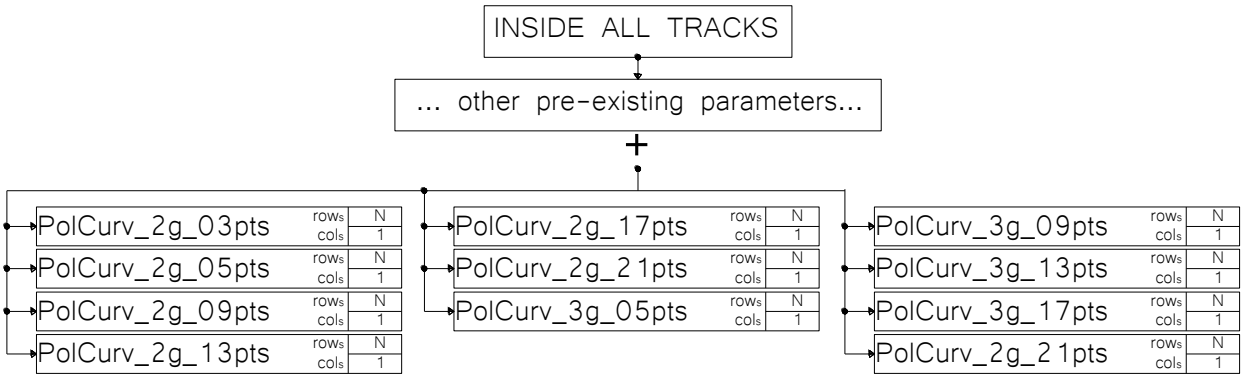


Figure 5.8 – Updated database detail after local polynomial curvature evaluation

5.2.4 Local curvature with axis methodology

Similarly to polynomial fitting, local curvature has been evaluated also with the axis methodology [2.3.2]. Each point of each track (except for the first and last point) has been associated with a local curvature value, based on the local circumference passing through three consecutive points. The curvature has been assigned, as usual, to the central point of the set, and has been stored inside the database in the variable CircCurv. The logical process of the algorithm is shown in Figure 5.10, while the algorithm itself is available in attachment: "*Local curvature with axis methodology*".

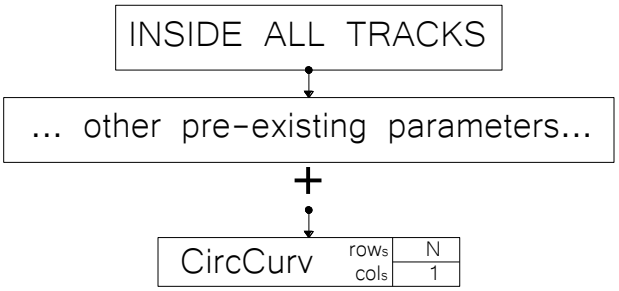


Figure 5.9 – Updated database detail after local curvature evaluation with axis methodology

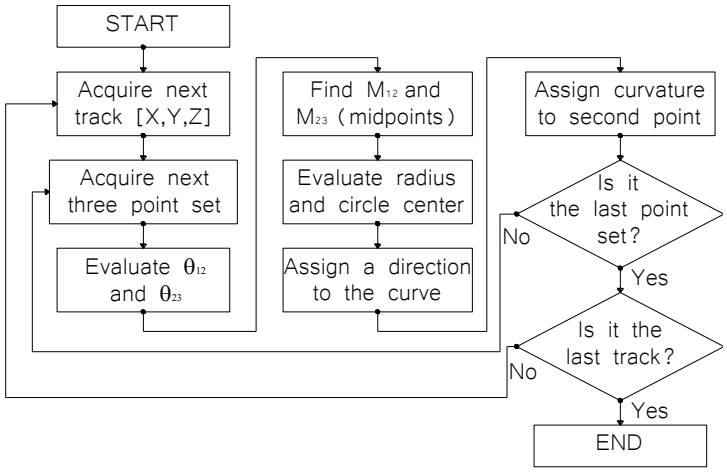


Figure 5.10 – Local curvature with axis methodology

5.2.5 Local curvature and deviation angle statistics

Since the database contains various and wide data and information, it has been decided to add some statistics to each tracks containing some fundamental information about the comparison between the real value of deviation angle and curvature and the measured one.
In order to make this comparison, two new vectors, REAL_DEV and REAL_CURV, containing the theoretical exact value of local curvature and deviation angle (extracted from design) have been manually added to the database.,In each single track some new variables have been stored:

- Errors_curv: curvature errors matrix, containing the difference between the local curvature calculated with all techniques and the exact theoretical value at each point;
- Errors_dev: deviation angle error vector, containing the difference between the local deviation angle calculated with the azimuth methodology and the exact theoretical value at each point;
- ErrParam_Curv: matrix containing average, standard error, maximum and minimum curvature error evaluated for each single techniques; and
- ErrParam_Curv: matrix containing average, standard error, maximum and minimum deviation angle error;

The algorithm itself is available in the attachment "*Local curvature and deviation angle statistics*", and it's not explained with a flow chart because it's only a simple comparison between exact and calculated values, while the new database setup is shown in Figure 5.11.

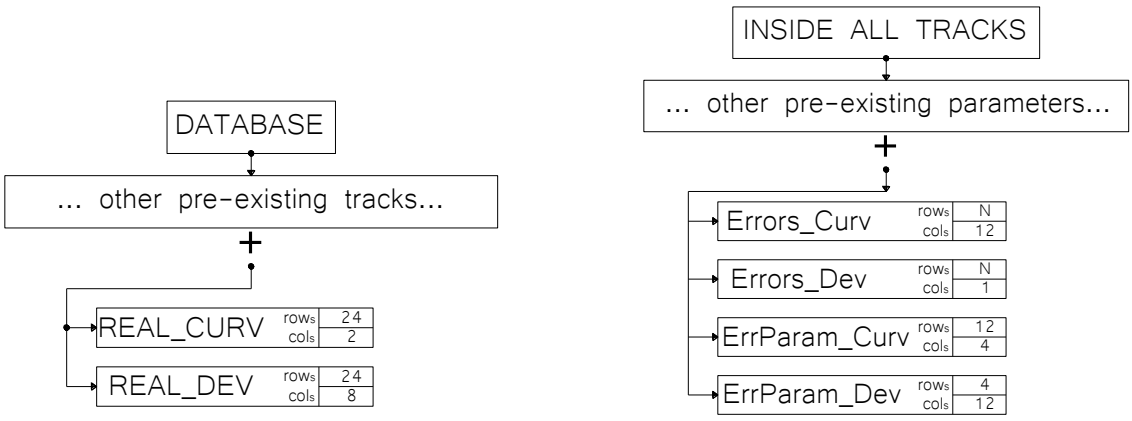


Figure 5.11 – Updated database detail after statistics analysis

5.2.6 Design-based grouping

During the work on the database some difficulties emerged to accomplish correctly the tracks grouping based upon local curvature and deviation angle evaluated with, respectively, polynomial fitting, axis methodology and azimuth differences. In order to evaluate fitting techniques [see 3.1] it has been decided to perform a design-based grouping, creating various points' subsets, one for each geometrical element of the alignment in each track. Consequently, inside each track it has been generated a new structure called **PROJ_SEGM** where all points' subsets for each geometrical element are contained. In addition, an initial variable called **ABSTRACT** has been added, which is a matrix of 25 rows and 7 columns, where there is the summary of the grouping:

- Each row of the matrix refers to a single geometrical element of the road (Table 4.1);
- Col #1 : Sequential number ID for each element;
- Col #2 : First point index position in track coordinates initial variable:
- Col #3 : Last point index position in track coordinates initial variable;
- Col #4 and #5 : First point coordinates (East and North); and
- Col #6 and #7 : Last point coordinates (East and North).

Tangency design points have been stored in the database inside a new variable called **GEOM_VERTEX**, containing coordinates of first and last point of each geometrical element of the track. The logical process of the algorithm is shown in Figure 5.12 and Figure 5.13, while the algorithm itself is available in attachment "*Design-based grouping*".

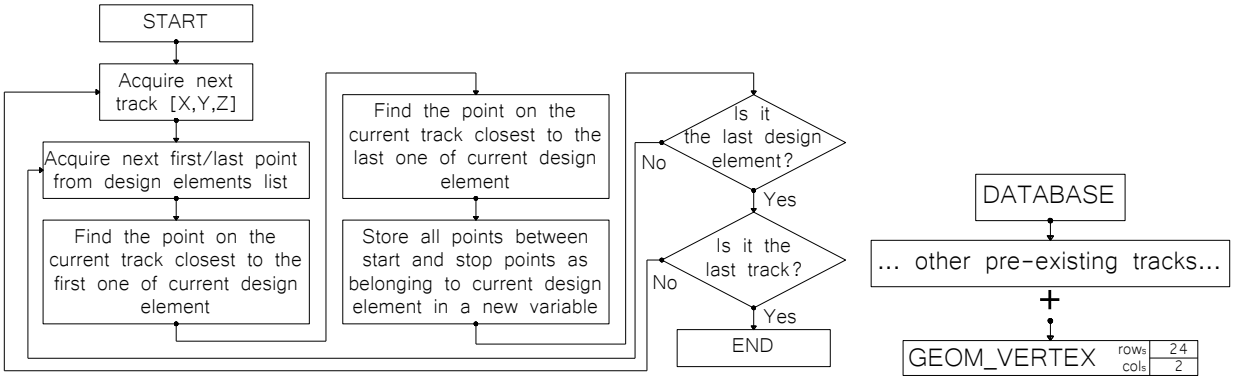


Figure 5.12 – Design-based grouping process and global database update

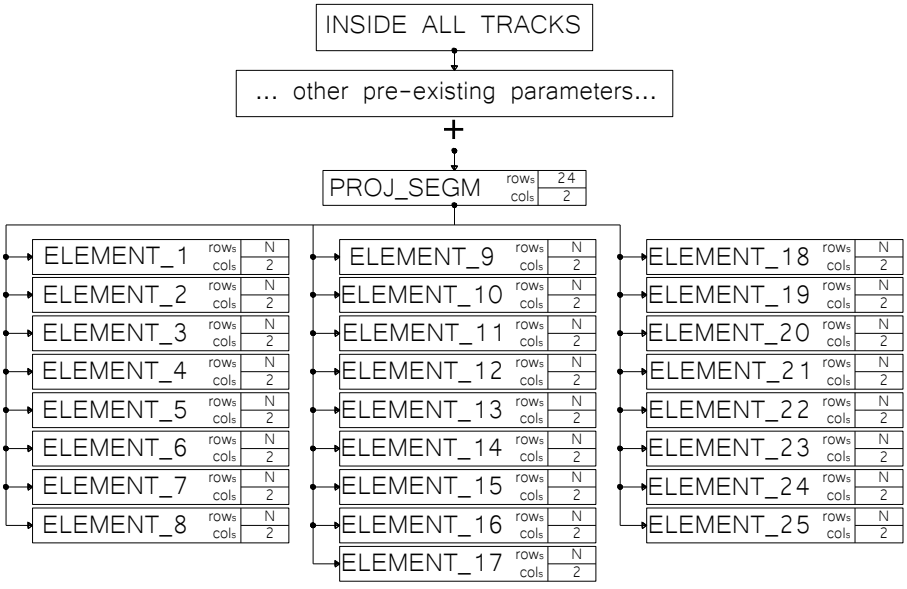


Figure 5.13 – Updated database detail after design-based grouping

5.2.7 Least squares, Huber and Landau fitting on design-based grouping

The design-based grouping gives the possibility to perform a complete fitting of each single geometrical element of each track with all the fitting techniques previously described. Performing this fitting represents an original way to verify the potential and application limits of each fitting algorithm without incurring in any error dependent on survey quality and consistency. This very specific methodology usually is not adopted in other works available in literature, where, instead, the fitting technique is created with the objective to fit the database. Furthermore, other usual working procedures generally lead to algorithms that are tipically able to work in limited applications and with a very high risk to misunderstandings in terms of results.
All fittings have been performed with all tracks and they have been stored inside the corresponding geometry element, which is organized through a dedicated variable. The logical process of the algorithm is shown in Figure 5.14, while the algorithm itself is available in the attachment "*Least squares, Landau and Huber fitting*".

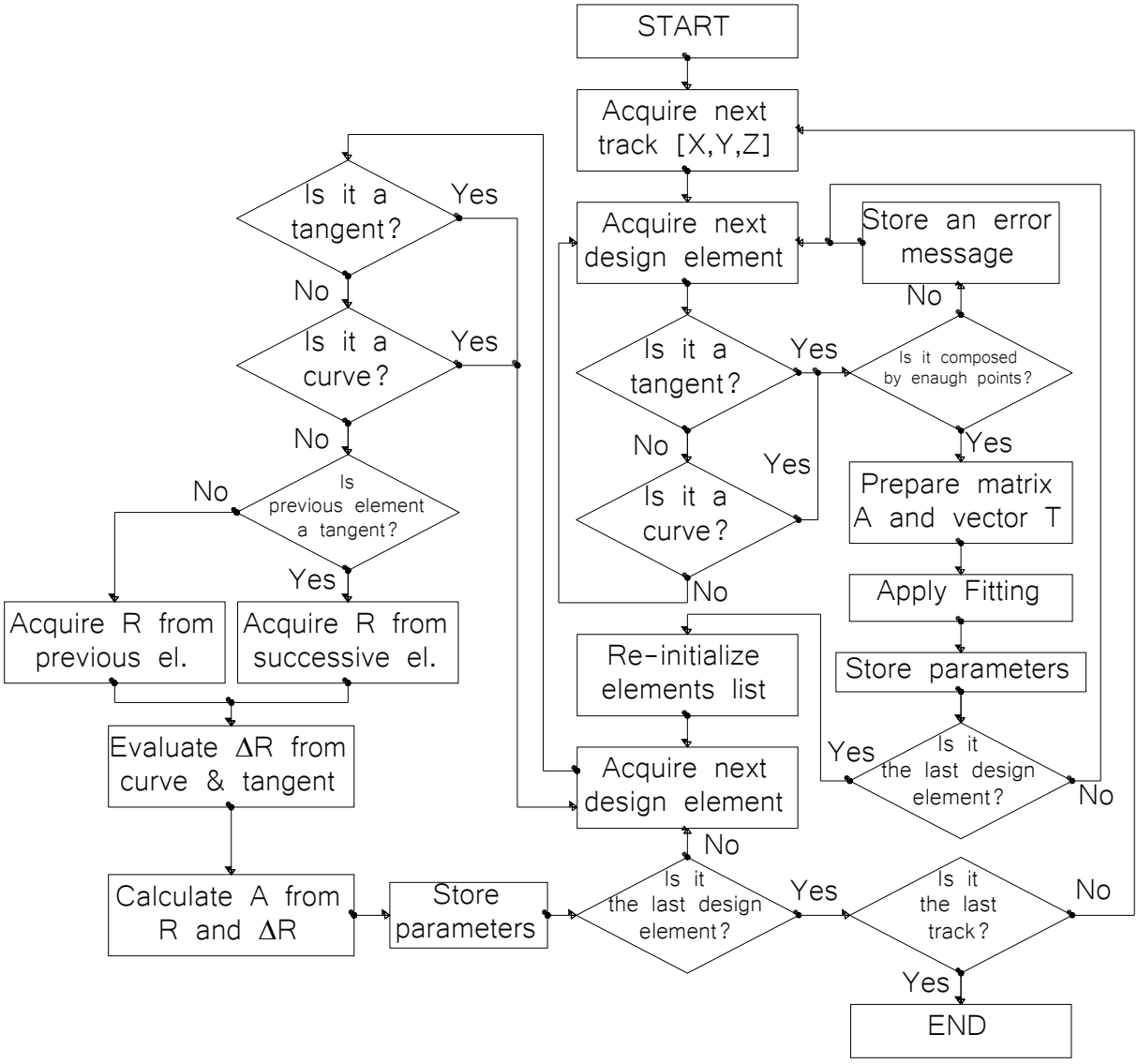


Figure 5.14 – Fitting algorithm process

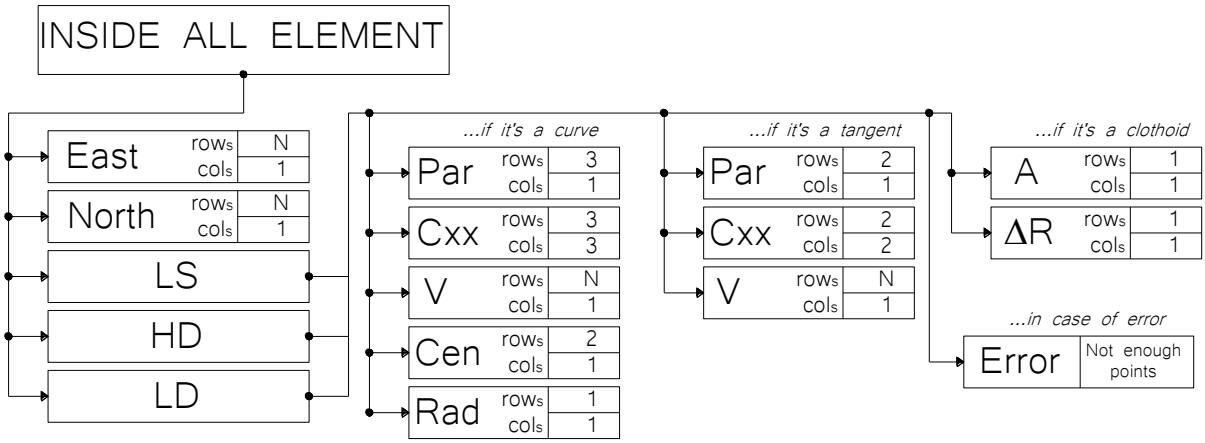


Figure 5.15 – Updated database detail after fitting

In addition to these detailed algorithms, other codes have been developed and implemented to manage the database and organize the big amount of data stored in it. Overall up to 40 scripts have been written and applied. Since it is not intended to give deep programming details inside this research work, but it is preferred to keep the focus on road-related topics, these specific scripts are not included in the current document, not even as attachment. Indeed, attaching many codes without any comment is relatively useless, nevertheless all codes and algorithms will be available on the CD attached to the current work, and they can be eventually distributed on demand to the author.

c6 Data processing and results

"The best results are obtained only with the maximum full passion"
Goethe

In this chapter, most meaningful graphs/charts/images related to the topics are shown and explained.

The amount of graphs and results that can be extracted from the database with provided algorithms and codes is really big, and it allows a deep study & analysis on all parameters investigated in this work, like local curvatures, local deviation angles and elements fitting. Inside the text will be inserted, as well as commented, at least one graph for each type, including those considered most significant and noteworthy.

6.1 GEOMETRY GROUPING

6.1.1 Local curvature

6.1.1.1 Axis methodology

In Figure 6.1 the behavior of the local curvature obtained through the axis methodology, or namely the construction of a circular curve perfectly fitting three consecutive points (epochs) on a track, is reported.
In the graph, the overall medium error on a single track (obtained as the average of all errors all along the track) on the y-axis, while on x-axis are displayed for various level of maximum random error artificially applied on tracks (from 2 cm to 40 cm); thereby, each single point on the graph represents a different track, characterized by a space-based frequency, and a maximum random error.
First, it is considered necessary to underline that the y-axis is in logarithmic scale, so that all errors, contained between 0 and 1, result as negative numbers. That is the reason why all the graphs are developed in the negative part of y-axis: so that it is intended that farther the point is from the x-axis, smaller is the error.

It is possible to state that there is a strong relationship between the local curvature error (obtained with local circumference) and both random maximum accuracy and space-based frequency. Looking to the details, it's clear that the effect of maximum random error on local curvature is almost negligible with respect to the effect given by the space-based frequency: the first one is able to reduce general accuracy remaining in the same magnitude order, while the only effect of space-based frequency can lead to big reduction of error, even up to three order of magnitude.
It is really interesting that there's a specific correlation between the medium error obtained on each track and the maximum random error applied, which is extraordinary well fitted with a 3rd degree polynomial function, with R^2-values between almost 0.96 and almost 1.

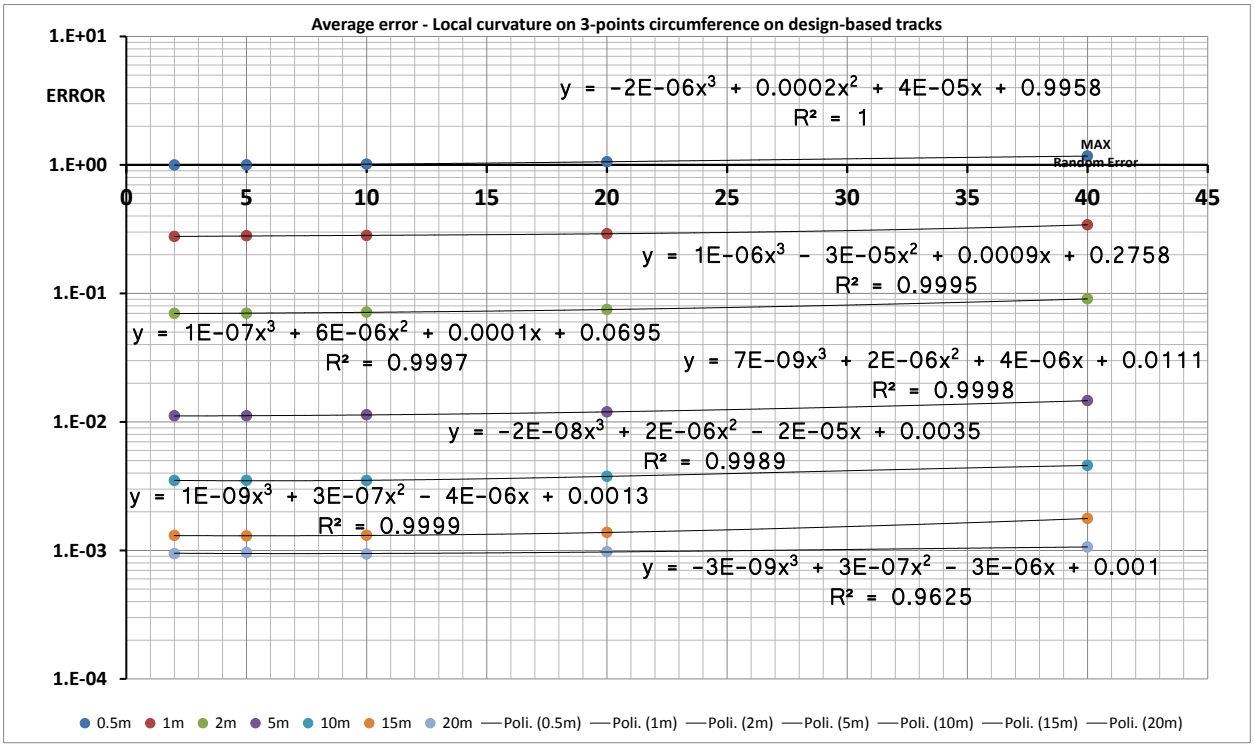


Figure 6.1 - Error - Local Curvature on 3pts circumference on design-based tracks

As an example of the behavior of local curvature depending on maximum random error and space-based frequency, in Figure 6.2 and Figure 6.3 two borderline situations are shown: the first one is the local curvature diagram with a space-based frequency equal to 0.5 m, while in the second It's equal to 20 m. These graphs are practically self-explained and show perfectly what is summarized in Figure 6.1: the effect of spacing in local curvature with axis methodology is definitely more strong and effective than the positioning point precision of the track.

In the foregoing graph, only tracks originated upon the design are shown and only for them it's possible to make a comparison with the theoretical design. This is because any inaccuracy on point's position leads to errors in terms of station and consequently some problems in terms of comparison between corresponding stations.
Practically, in order to make a comparison between the theoretical value and the calculated one it is necessary to have as reference the same station: on a surveyed track the only way to evaluate the station at each epoch is to calculate the sum of all the consecutive straight line segments up to the interested point; any error in positioning leads to overestimate distances and produces a shifting between the real station and the calculated one.
This effect is clearly recognizable in Figure 6.2: looking at the right part of the graph the curvature evaluated on tracks with high maximum random error (40 cm, 20 cm and 10 cm) seems to be longer than the others, up to a final error equal to 500 m. Conversely, where the random error is almost negligible respect to the space-based frequency (20 m in Figure 6.3) there is no shifting effect.

To fix this spatial synchronization it is possible to calculate a "surrogate" station knowing the theoretical spacing between each single epoch of the track. If the tracks have been sampled with a space-based frequency, it's possible to remedy to positioning error by evaluating the station with a progressive sum of the spacing adopted (e.g., if the track has a space frequency equal to 1m, the first epoch/point will have station zero, the second point will have station 1 m, the third point will have station 2 m, etc...).
The knowledge of the way each track has been generated explain that this specific procedure is only possible if tracks are characterized by a **fixed** space-based frequency between each single epoch. In this work, this assumption is valid only with design-based tracks, but not with surveyed tracks, which have been obtained by the image interpretation, with a variable speed and with a time-based frequency (50Hz on IMU and 1Hz on GPS).

Accordingly, only design-based tracks have been characterized by an error estimate in terms of curvature and deviation angle, while with surveyed tracks it's only possible a qualitative analysis on plotted graphs.

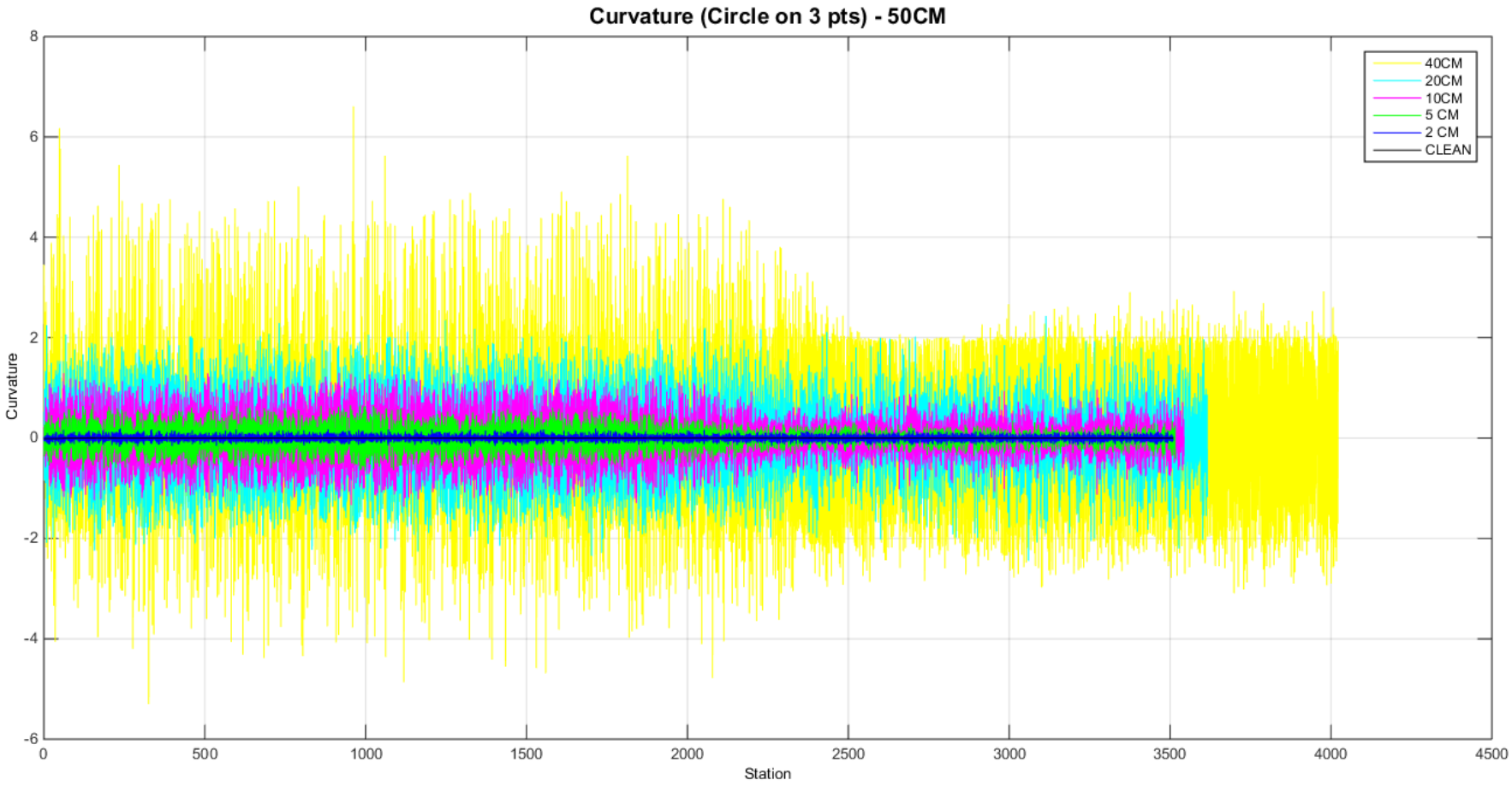


Figure 6.2 - Local circumference curvature (3 pts) - design-based tracks – 50cm spacing

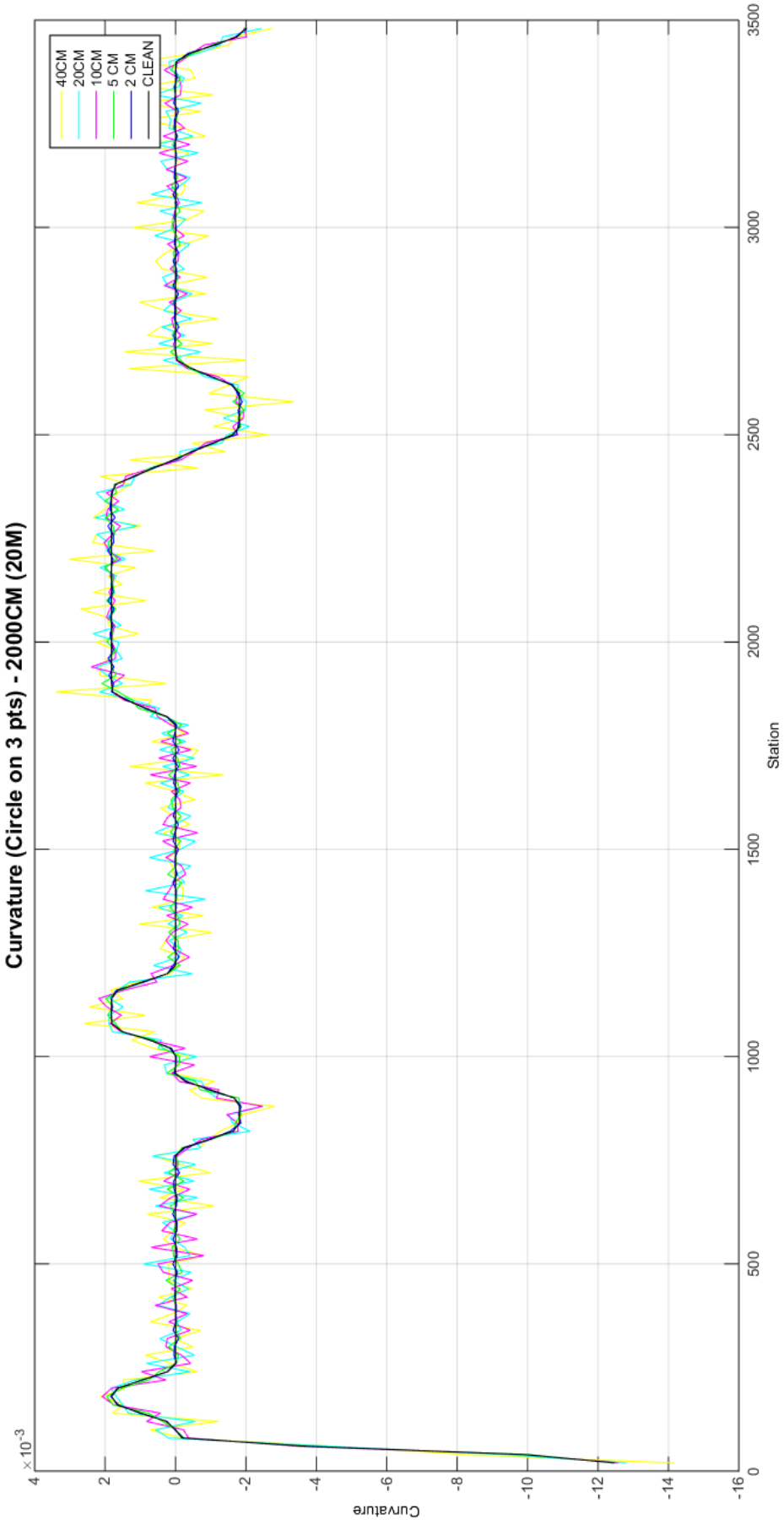


Figure 6.3 - Local circumference curvature (3 pts) - design-based tracks - 20m spacing

Surveyed tracks, instead, do not give the possibility to recognize any of these regular relationships or correlation, but the quality of the curvature obtained is really poor and generally does not lead to any possible recognition of geometric elements of the track. Unfortunately, analyzing surveyed tracks, and especially those from image analysis, only random noise is detected, with absolute no correlation with any real value: this behavior is present in

- Round trip vehicle trajectory average (IMU_Traj)
- Round trip roadside average (IMU_GigLine)
- Round trip centerline average (IMU_CntLine)
- Round trip lateral line average (GPS_LatLine)
- Round trip roadside average (GPS_CigLine)
- Round trip centerline average (GPS_CntLine)

An example of this noise detection is available in Figure 6.4: the yellow line represents the local curvature as it has been calculated with the 3 points local circumference. It should approximate the black line (the design curvature) as best as possible, but it actually represents a useless noise, that cannot be corrected even with a filter (e.g., like a moving average). In addition, in this case, there is a huge error in terms of station: at the end of the track is registered a shifting up to 600m.

Better results have been obtained with the other surveyed tracks, where the curvature is still noisy but it gives reasonable values and it can be processed with filters. Some examples are visible in the following graphs (Figure 6.5 and Figure 6.6).

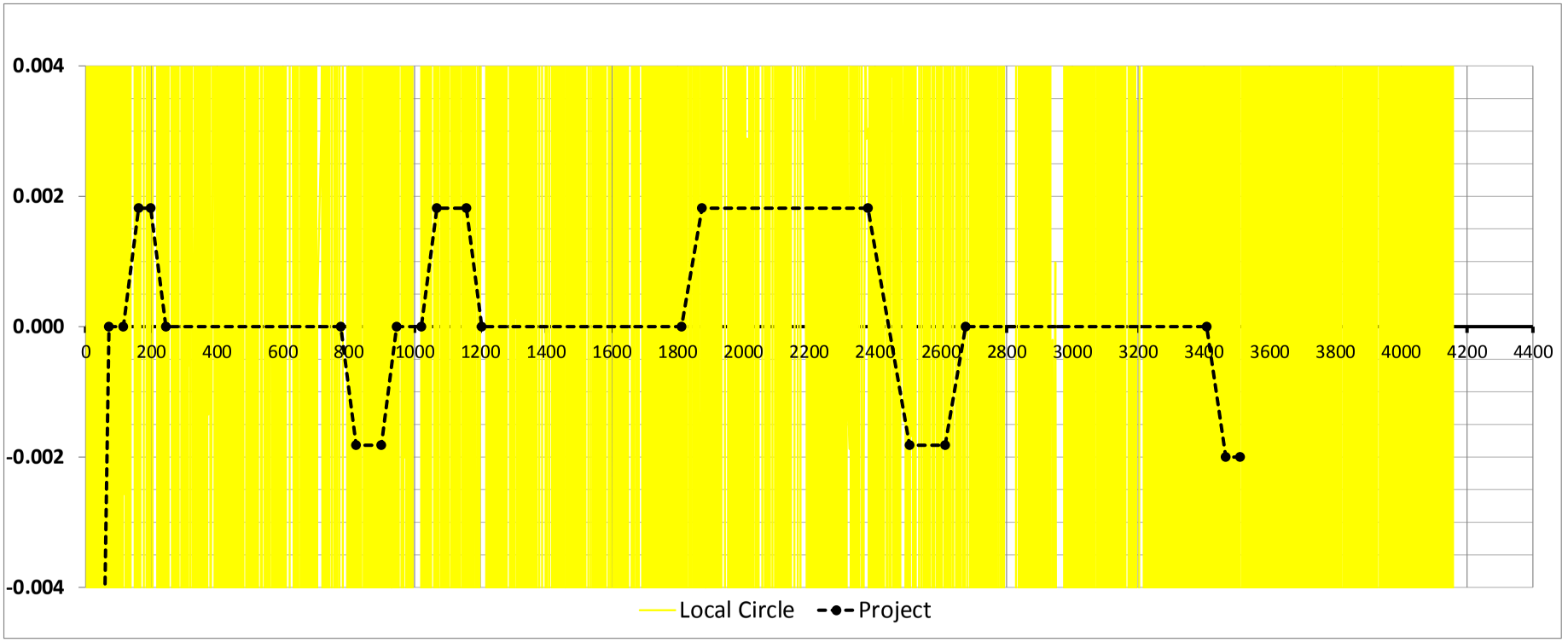


Figure 6.4 - Local circumf. curvature (3 pts) – survey-based tracks with high random error (GPS_CigLine)

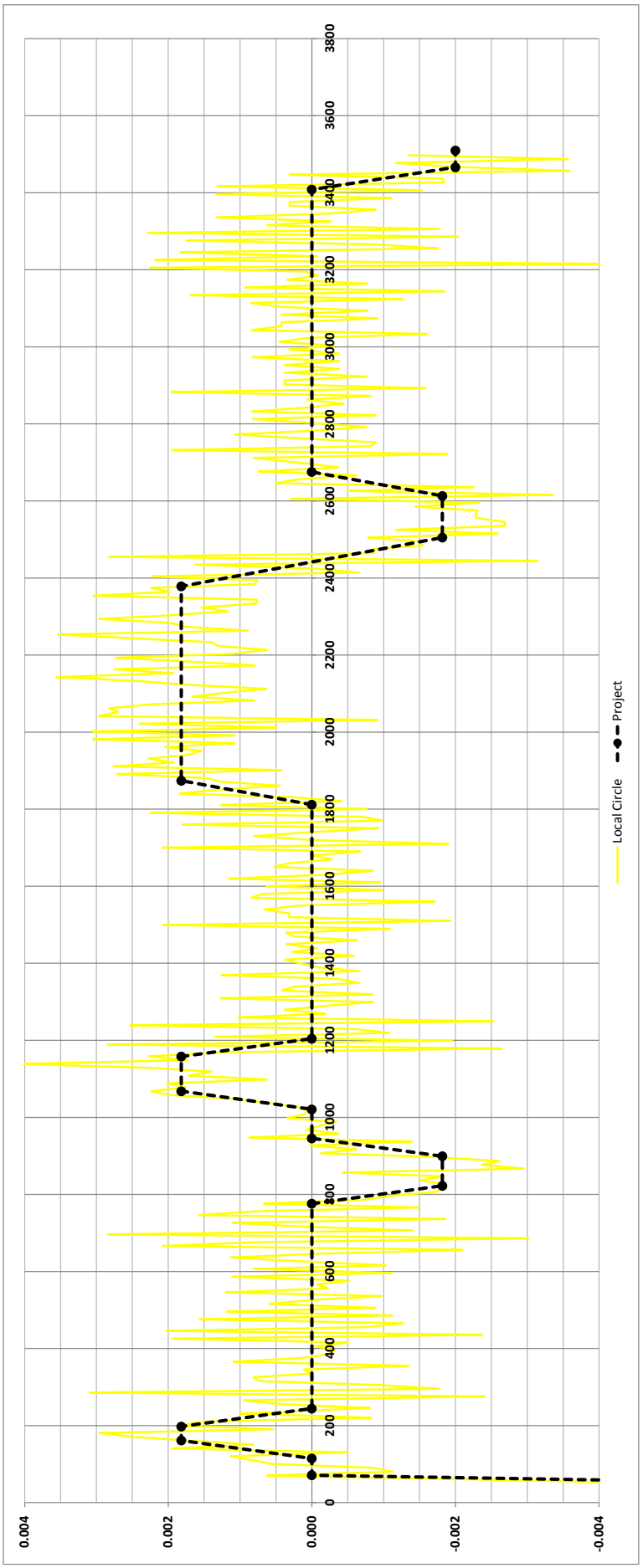


Figure 6.5 – Local circumference curvature (3 pts) – survey-based tracks (GIS CntLine)

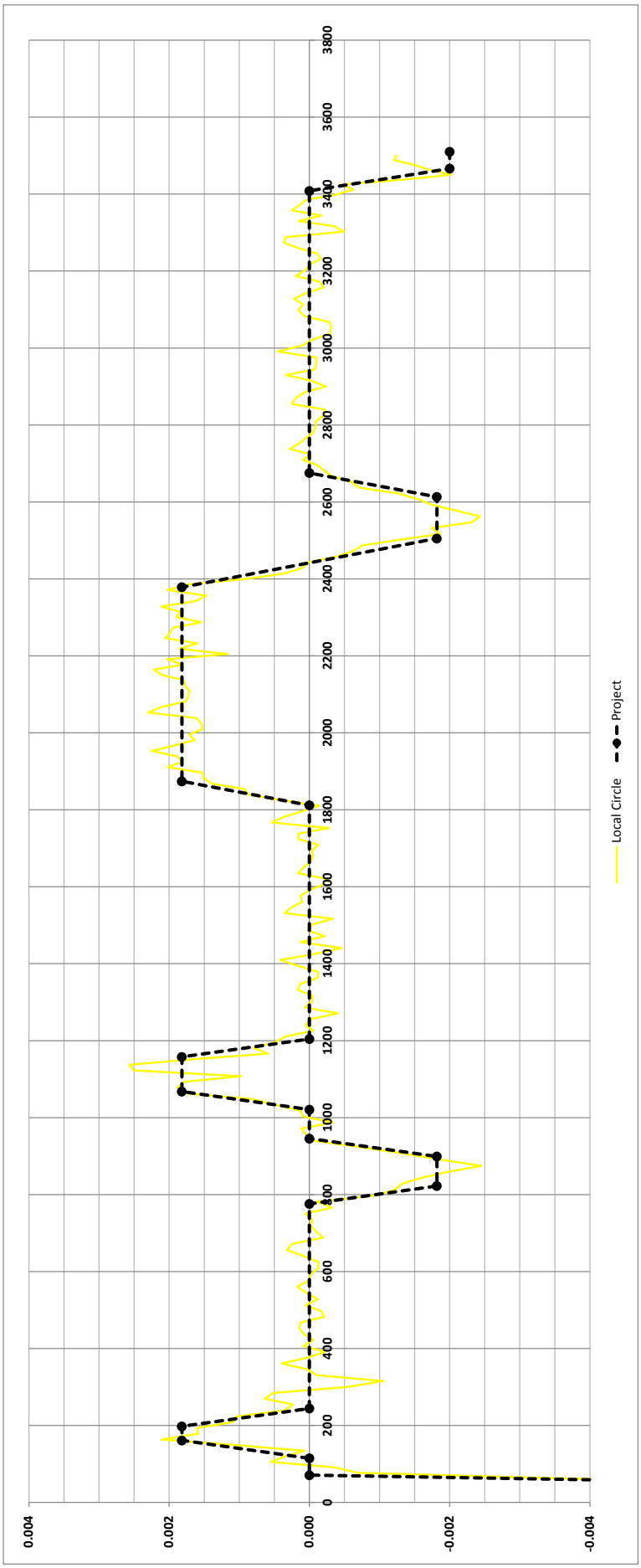


Figure 6.6 – Local circumference curvature (3 pts) – survey-based tracks (GPS-Traj)

6.1.1.2 Polynomial fitting

Polynomial fitting generally gives back results and values that are really effective (if compared to the local circle) thanks to its ability to control local discontinuities. Due to its "fitting" nature, that is not a characteristic of the local circumference, it is possible to avoid sharp and sudden oscillations, like those one present in previous graphs.

From Figure 6.7 up to Figure 6.13 there is the overall statistical analysis of curvature obtained with polynomial fitting. Each single graph is referred to a predetermined space-based frequency (from 0.5 m to 20 m): on Y-axis it is shown the medium error (obtained as the average of all errors all along each single track), while on the X-axis various width of polynomial fitting mobile windows are displayed. As explained in (2.3.3), polynomial function is a fitting techniques that is able to investigate detailed areas and also really wide areas, and this is possible thanks to the fitting windows width; for each single fitting it's possible to set how many points have to be included in the fitting process around the interested point/epoch: the wider is the windows, the larger is the fitting area and smoother is the results.
In this work, it has been decided to investigate different fitting windows width (3, 5, 9, 13, 17 and 21 points) and they are composed by an odd number of points to univocally individuate the central point to which the curvature will be associated. For each single maximum random error and space-based frequency, the track has been fitted with two different polynomial degrees (2^{nd} and 3^{rd}) and, with all windows width, obtaining up to 11 different curvatures: in each graph each column represent a single combination of fitting parameters (polynomial degree, maximum random error, and windows width).
In addition, in these graphs, the Y-axis is in logarithmic scale and all errors result as negative numbers. It is intended that farther the point is from the x-axis, smaller is the error.

In all figures, it is possible to recognize a progressive decreasing in medium error together with a growing fitting window width: generally, it is possible to state that much wider is the analysis (far from detailed area investigations) smaller are the errors. The effect of fitting window width is relevant and, like the curvature measured with local circle, it's able to reduce the average error up to 4 order of magnitude.
On tracks with no random error applied (which are exactly located on the design original alignment) the curvature error is strongly reduced, down to level close to machine accuracy, and it keeps generally close to zero; the only exceptions are visible in Figure 6.12 and Figure 6.13, where there is a small increase in error. This last phenomenon is attributable to the excessive fitting window width that induces instability in the fitted polynomial function: in extreme parameters combinations the window can be up to 400 meters wide and this measure is definitely not suitable to perform an accurate and detailed curvature evaluation.
Looking at error-affected tracks, it is possible to state that there is no sensitive difference related to polynomial degree and maximum random error: the results are definitely of the same magnitude and, like with no-error affected tracks, even in this case the average error generally increases together with the increasing of the fitting windows width.

The general error tendency is to increase with the spacing and the fitting window width, and this last one can lead to big variation in error.

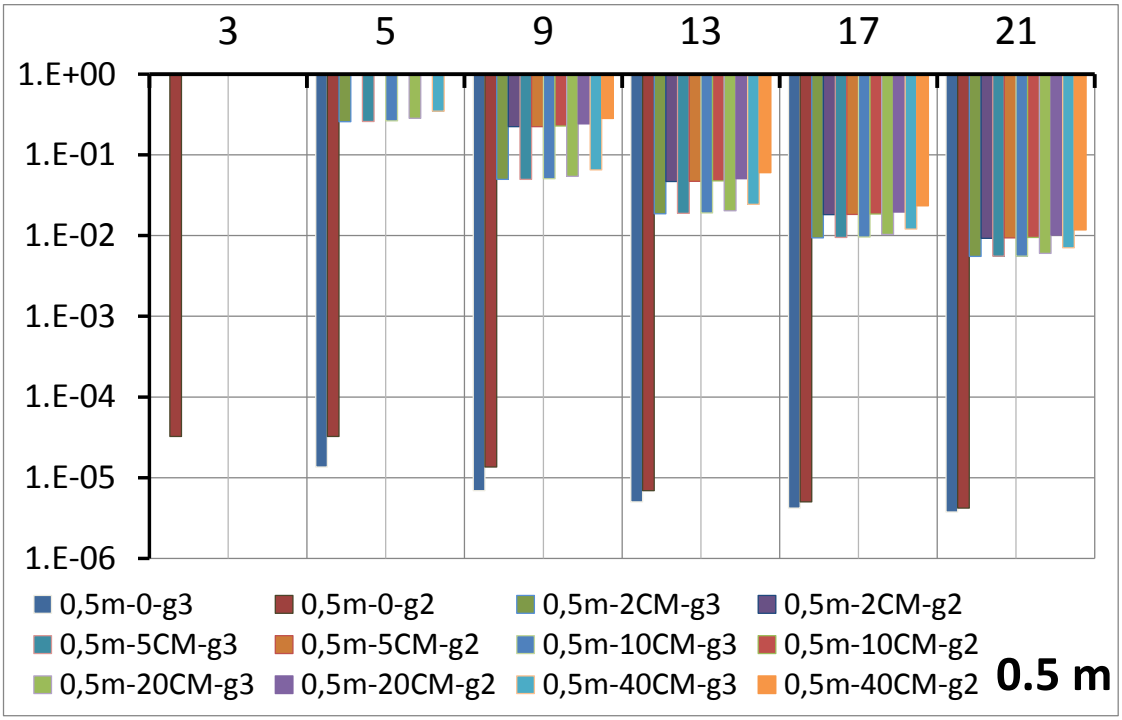


Figure 6.7 – Polynomial curvature in 0.5m design-based tracks

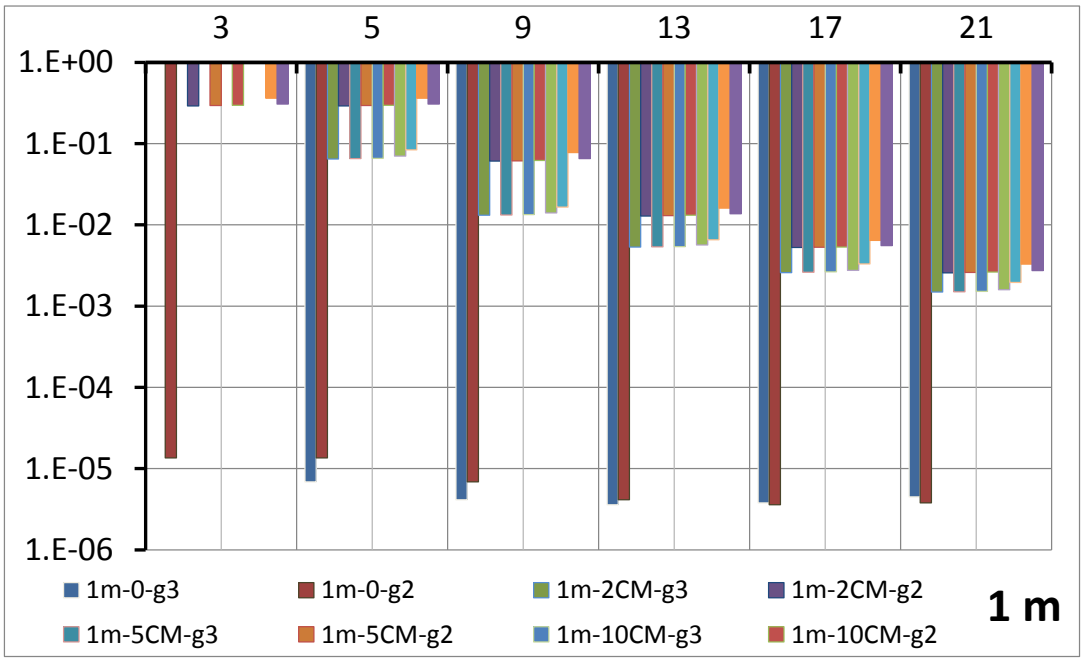


Figure 6.8 – Polynomial curvature in 1m design-based tracks

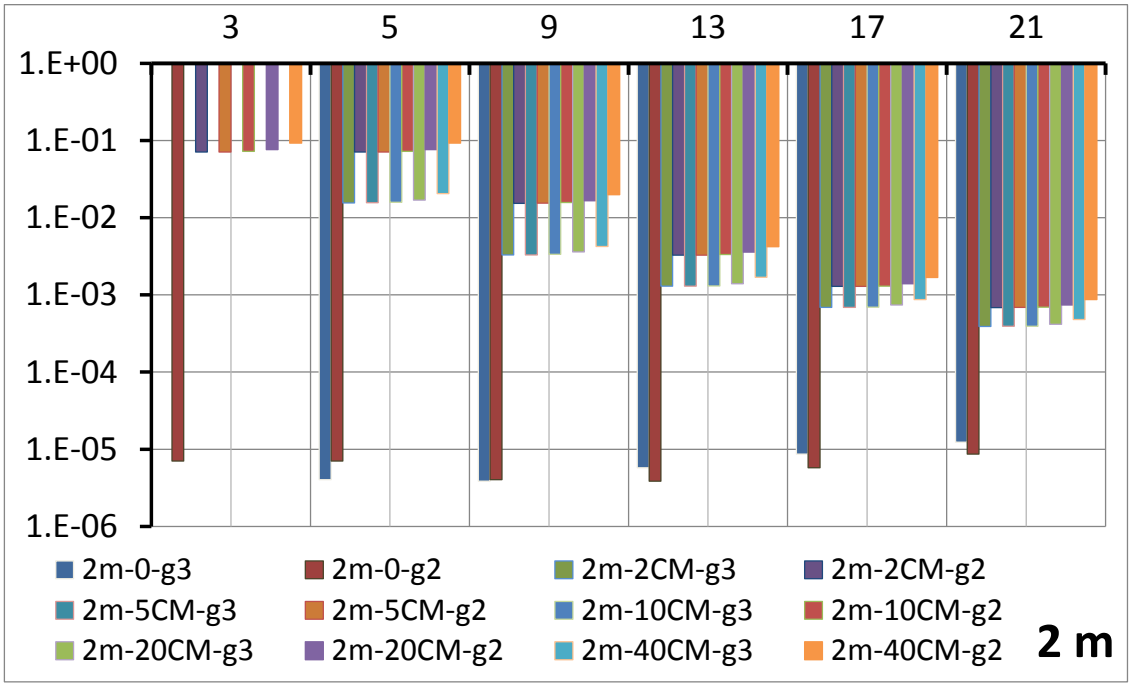


Figure 6.9 - Polynomial curvature in 2m design-based tracks

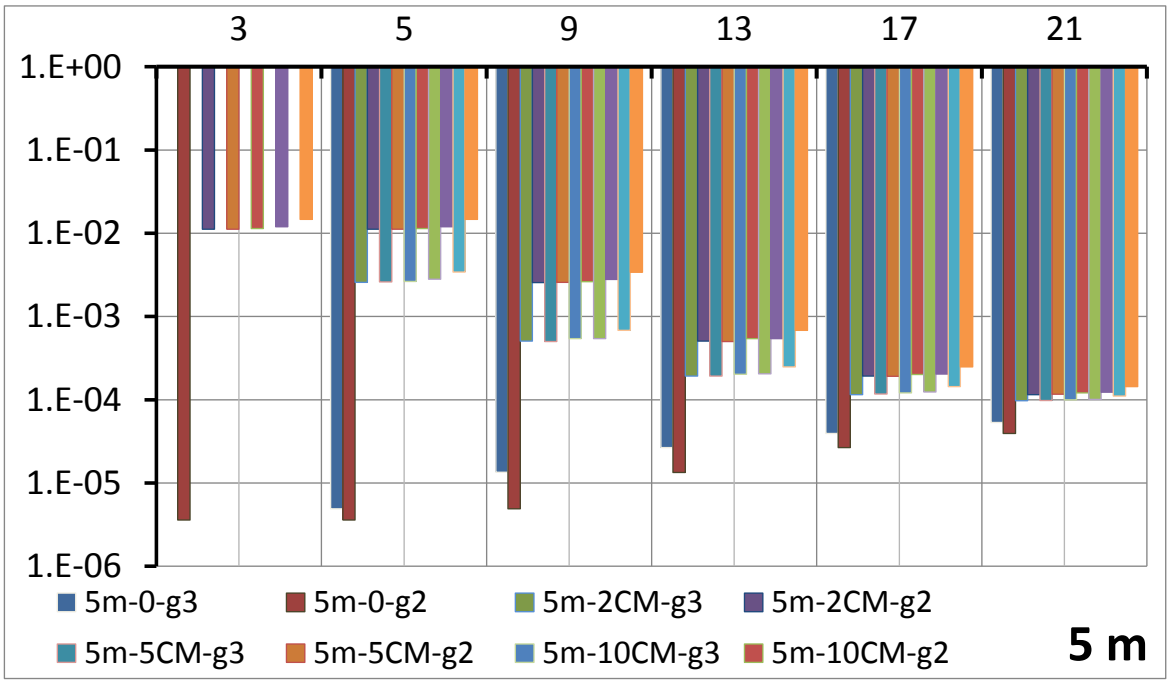


Figure 6.10 - Polynomial curvature in 5m design-based tracks

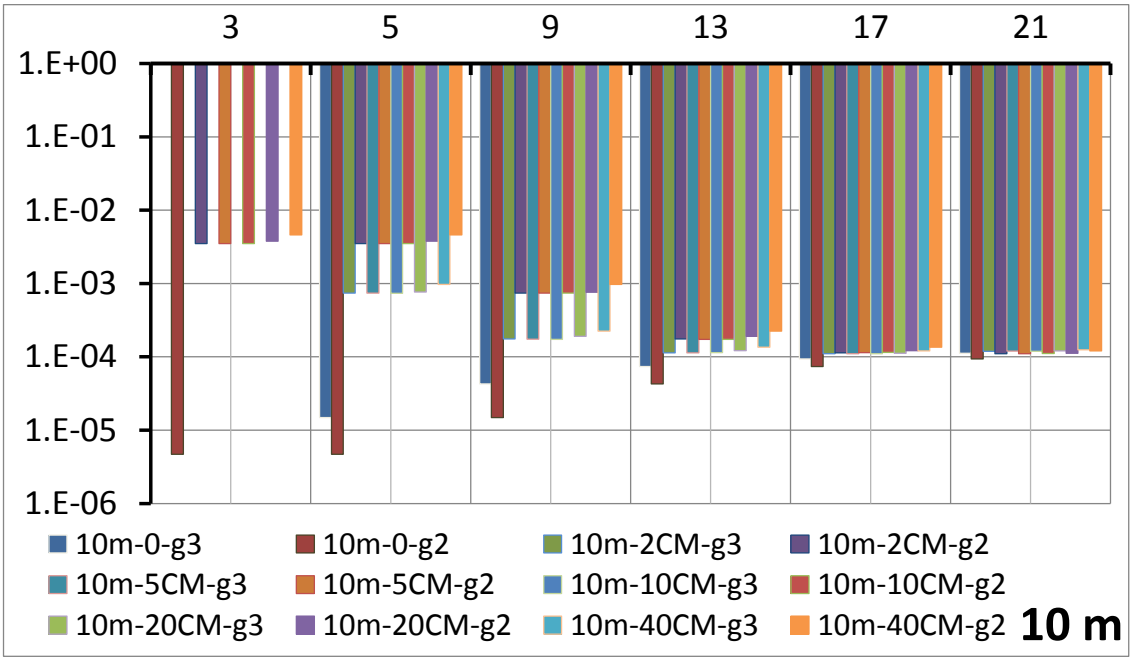


Figure 6.11 - Polynomial curvature in 10m design-based tracks

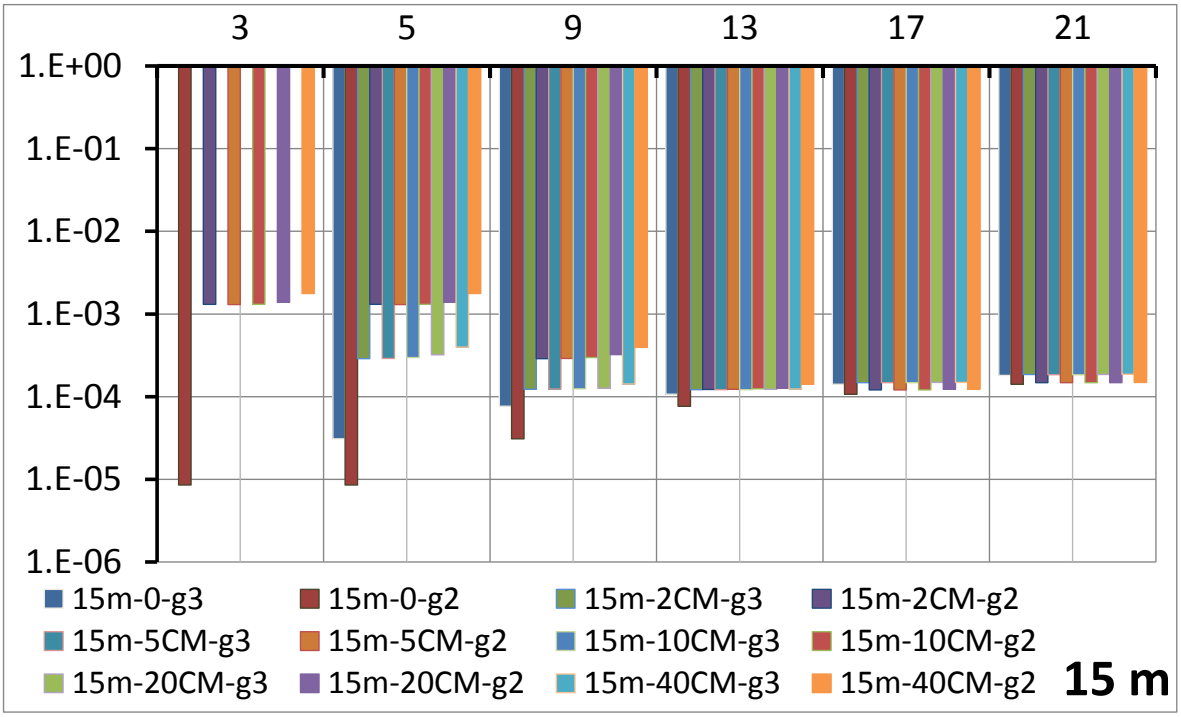


Figure 6.12 - Polynomial curvature in 15m design-based tracks

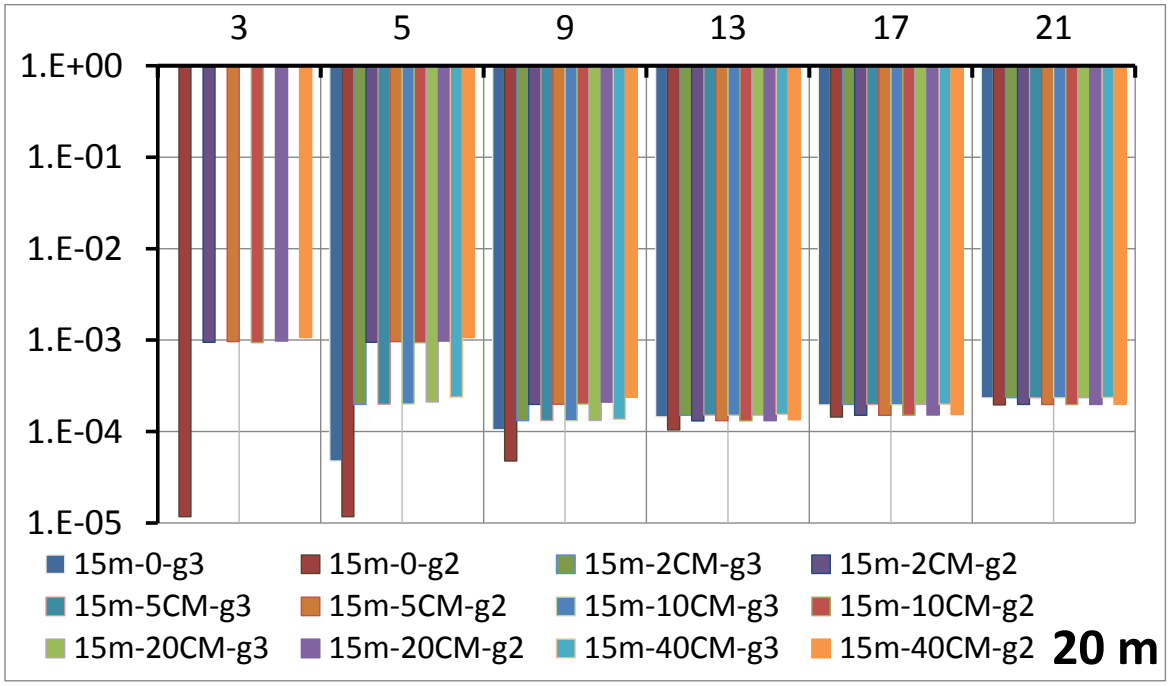


Figure 6.13 – Polynomial curvature in 20m design-based tracks

Like with the axis methodology, it has not been possible to make an accurate comparative statistics with real values, due to the already mentioned station shifting that affects this kind of tracks. Even in this case, a qualitative comparison has been made.

As an example, some curvature diagrams are visible in the following graphs.

Definitely, the best solution in terms of accuracy and signal regularity is the one available in Figure 6.15, referred to the local curvature obtained on the GIS centerline manual sampling. Actually, this result is quite surprising, because theoretically the GIS aerial picture is the less accurate data origin on which it has been worked on. The accuracy available from the picture is a value around 40 cm, and this specific threshold is generally based upon the pixel definition of the image. In Figure 6.14 it is possible to see a detailed zoom in the aerial image, and to notice that white lines on the carriageway are not uniquely determined, but they are recognizable by the lighter pixels area on the picture.

So the point is that a really inaccurate source leads unexpectedly to really good results.

This apparently strange result is the proof that the human influence is still able to compensate theoretical inaccuracy and leading to good results. Actually, even if the picture quality is low, the human eye and brain can recognize the correct position of the white lines and it is able to face with the low resolution. Looking carefully at Figure 6.14, the sampling of the white line is quite regular and isn't affected by the confused pixel web below: the human brain can easily understand where is the center of the white line even if the "white pixel area" is quite confused or not enough restricted. Just picking up the central point of the area, the inaccuracy of the picture is strongly reduced, giving back a substantially regular and well accurate point set.

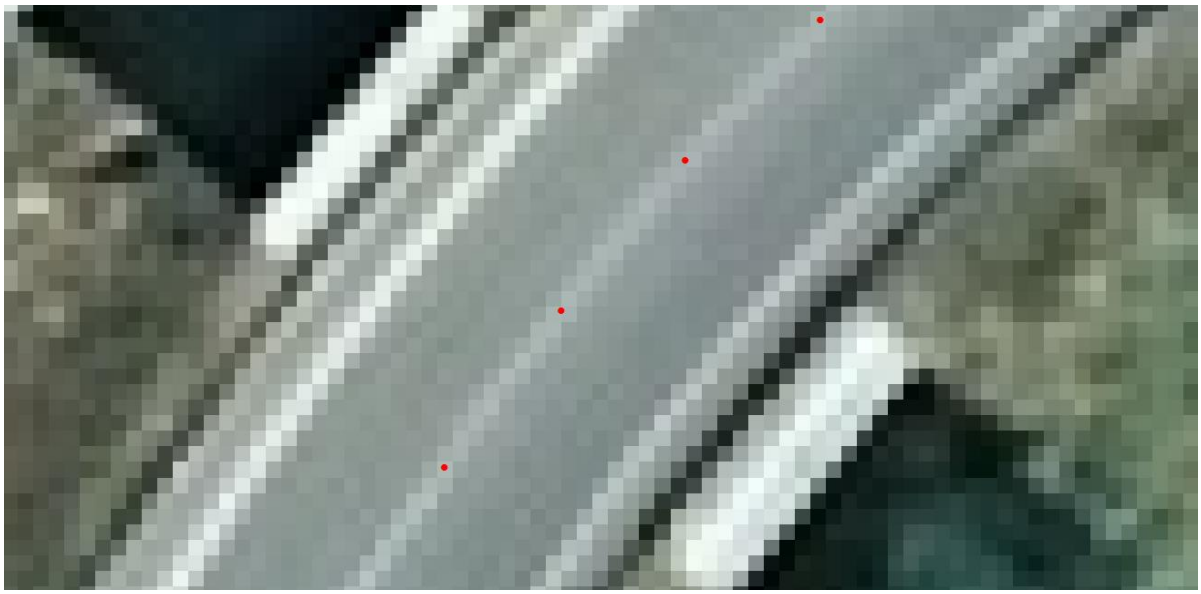


Figure 6.14 - Aerial GIS image detailed zoom

In Figure 6.17, the 3rd degree local polynomial curvature obtained with the average vehicles trajectories based on a GPS positioning reference is visible. This solution is quite satisfactory and it is impressive how effective is the change in the fitting window in terms of result. Widening the fitting window generates an effect similar to a moving average, which gives more precise results but with a lower resolution. In fact, the wider is the window, the smoother is the signal: this means that when it's not necessary to get detailed information but the target is to recognize general elements a wider fitting window can be adopted. Typically, to recognize geometrical elements boundaries (start and stop stations) could be useful to adopt a wide fitting window and being able to detect specific peaks, while to recognize small elements and characteristics the fitting window must be narrowed.
Worst results are obtained on IMU-based positioned tracks. IMU positioning accuracy is far smaller than the GPS one, so it is quite predictable that image analysis and curvature extraction from these set of tracks (IMU based) is less accurate and effective. Surely, the best solution is obtained with the trajectory average (Figure 6.16) where medium acceptable results are obtained with a fitting window width greater than 13 points. Despite everything, best signal quality (fitting window width equal to 21 points) is quite satisfactory and with the application of a moving average can possibly leads to peaks identification.
Due to the positioning low quality, a little station shifting is present, giving errors in the total length of the track and avoiding an accurate statistics.

The last graph shows the results obtained with the mobile mapping image analysis Figure 6.18. Actually, these sets of tracks do not show good performances neither with local curvature nor with deviation angles. This is surely due to big error propagation: starting from positioning accuracy (that in some cases, like with the IMU, it is small) and ending to image alignment recognition and geo-referencing, errors increase in each step, damaging the quality of any successive analysis. There is again a huge random noise instead of a useful signal and the station shifting is high.

This is the overall results summary in terms of local curvature evaluated with polynomial fitting and local circumference on 3 points. A complete and wide plethora of diagrams and graphs can be produced from the database and analyzed one by one. In order to have a look at most meaningful graphs is necessary to refer to attachments *Curvature profiles: axis methodology* and *Curvature profiles: polynomial fitting.*

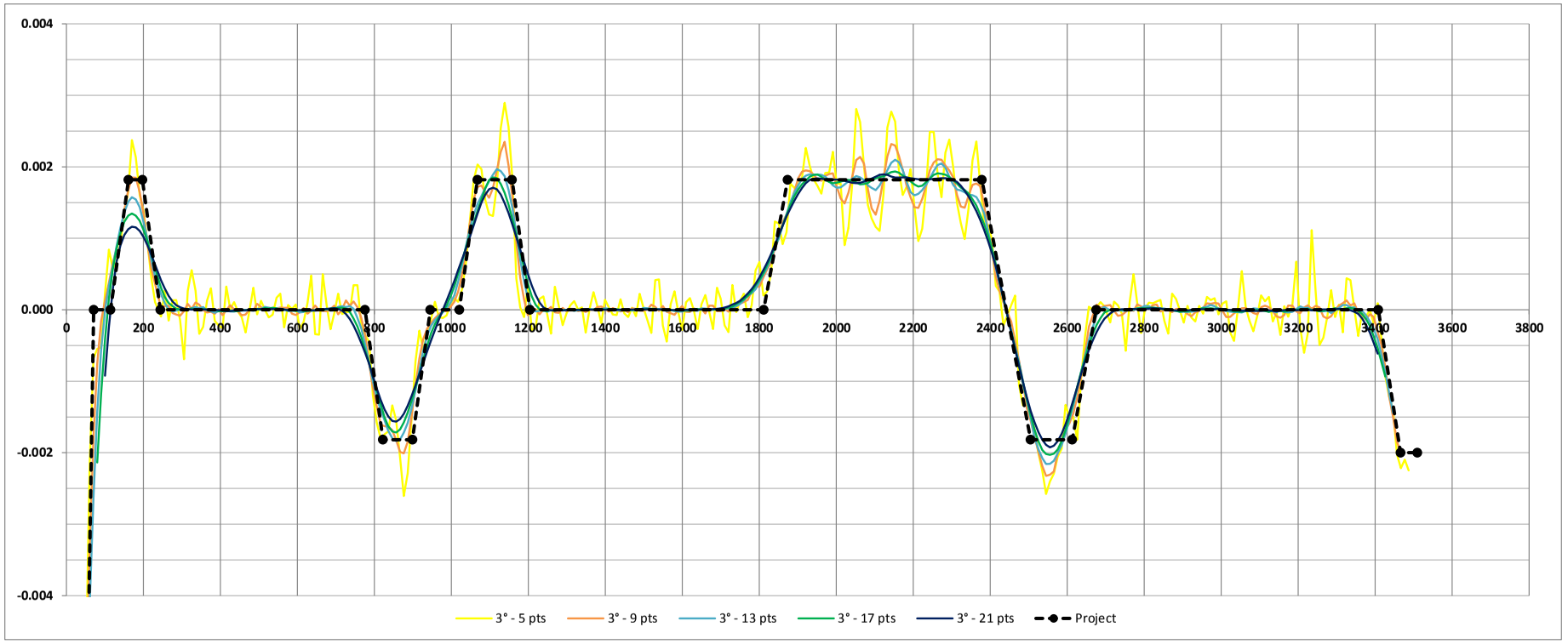


Figure 6.15 - Polynomial curvature (3° degree) – survey-based tracks (GIS-Centerline)

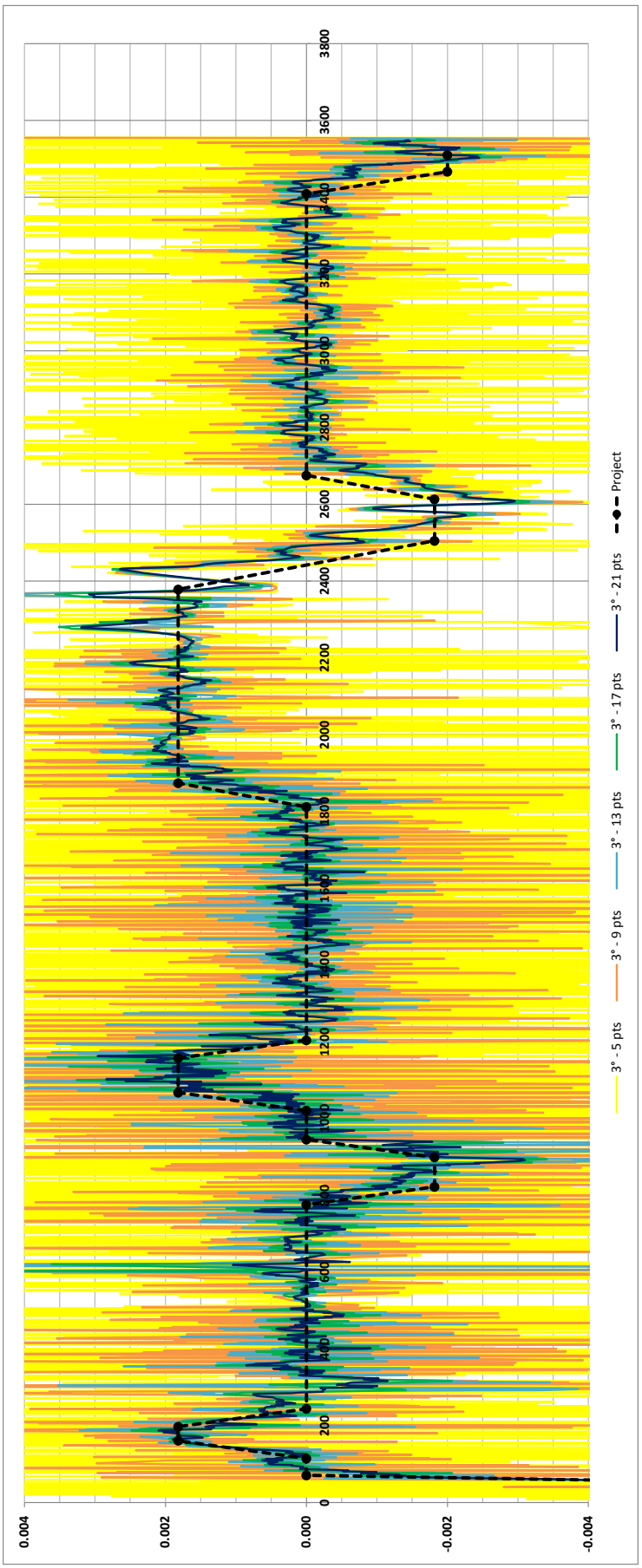


Figure 6.16 – Polynomial curvature (3° degree) – survey-based tracks (IMU-Trajectory)

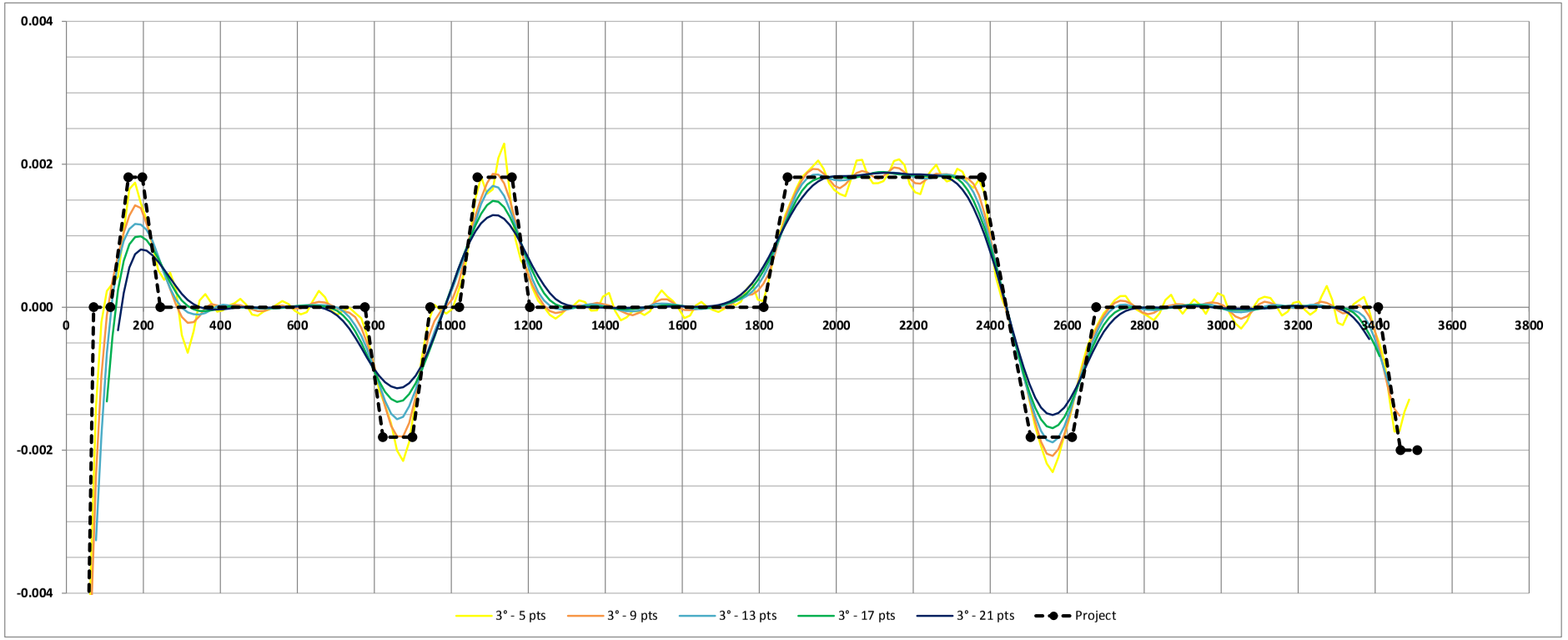


Figure 6.17 - Polynomial curvature (3° degree) - survey-based tracks (GPS-Trajectory)

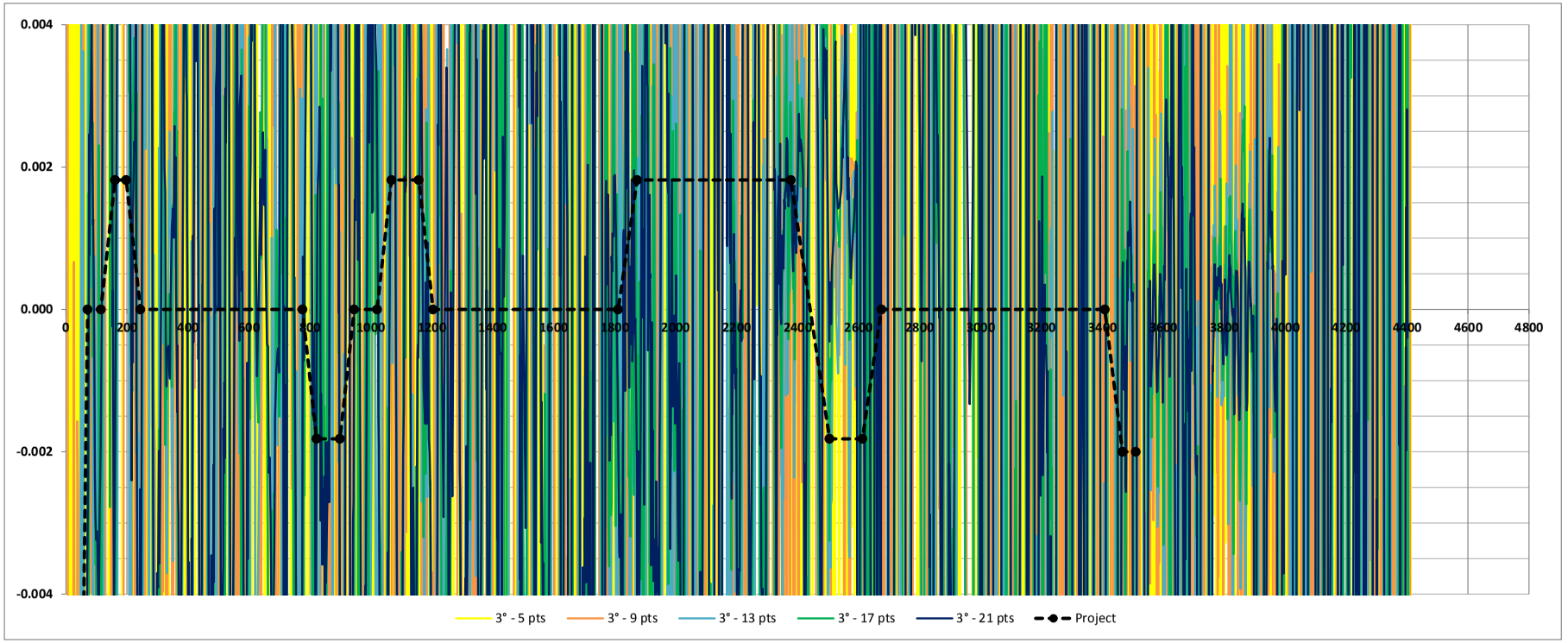


Figure 6.18 - Polynomial curvature (3° degree) – survey-based tracks (IMU-CigLine)

6.1.2 Local deviation angle

Figure 6.19 gives an overview of deviation angle medium errors vs the maximum random error applied. On Y-axis is shown the medium error (obtained as the average of all errors all along each single track), while on the X-axis are displayed the various width of polynomial fitting mobile windows.

This picture summarizes the quality of the deviation angle extraction and gives the possibility to recognize some impressive regular behavior of errors. A specific function relates medium accuracy and maximum random error applied: it has been recognized in a 3rd degree polynomial function that is able to fit each single curve almost perfectly, with R^2 going from 0.962 to almost 1. The medium error increases with the third power as long as the maximum error increase, or in other words, the accuracy of deviation angle extraction increases as long as the accuracy of the database increase, with a 3rd power relation. Furthermore, the medium error is heavily affected by the space-based frequency of the track: the smaller is the frequency, the higher is the error.

This last relationship is one of the most relevant, because it gives important references useful in terms of survey planning: the choice of a low space-based frequency (or superimposing an high speed to a Mobile Mapping vehicle with time-based acquisition frequency) is surely a good way to avoid high errors in deviation angles.
Furthermore, the bigger is the frequency, the bigger is the theoretical deviation angle and the smaller the error is, and these two elements both concur to increase the quality of the survey and the stability and accuracy of evaluated deviation angles.

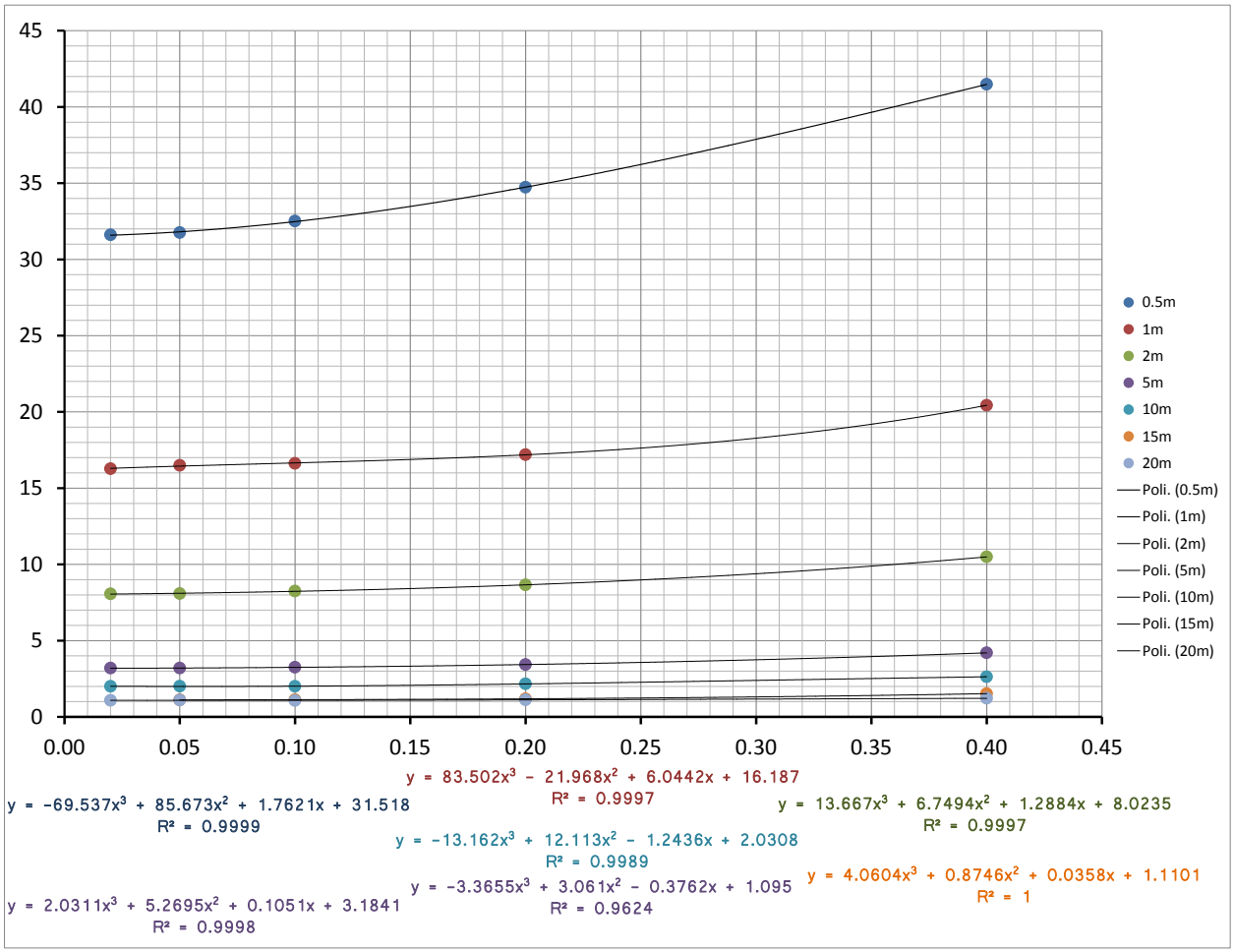


Figure 6.19 - Average error - Local deviation angle on design-based tracks

Figure 6.21 shows the extended deviation angle obtained from the design-based track with a space-based frequency equal to 50 cm and various level of maximum random error. It is easy to notice that it is not possible to recognize any path or element, but the deviation angle is actually a random noise. To have a better idea of the magnitude of the noise, in Figure 6.20 it has been applied a stretch on the Y-axis, making a zoom in the area with values between -5 and 3 degrees. It is visible, in black, the deviation angle obtained from the "clean" track (that is the original track without any random error), while any other deviation angle (from 2 cm to 40 cm of random error) is basically a random noise with absolutely no relation with the correct angle. There is no necessity to underline again the station shifting phenomena.

Conversely, Figure 6.22 shows the best solution obtained from the bigger space-based frequency tracks (20 m). In these tracks, the deviation angle is definitely good and can lead to good road geometrical elements detection. The maximum random error plays a key role and when it is particularly high it produces a noise that can leads to some problems and instabilities; station shifting is absent.

Looking at the deviation angle of the survey-based tracks no good results are obtained, except from the most accurate tracks: GIS_Centerline and GPS_Trajectory (Figure 6.23). In these tracks, it is possible to recognize some trends and curves are *visibly* distinguishable from tangents.

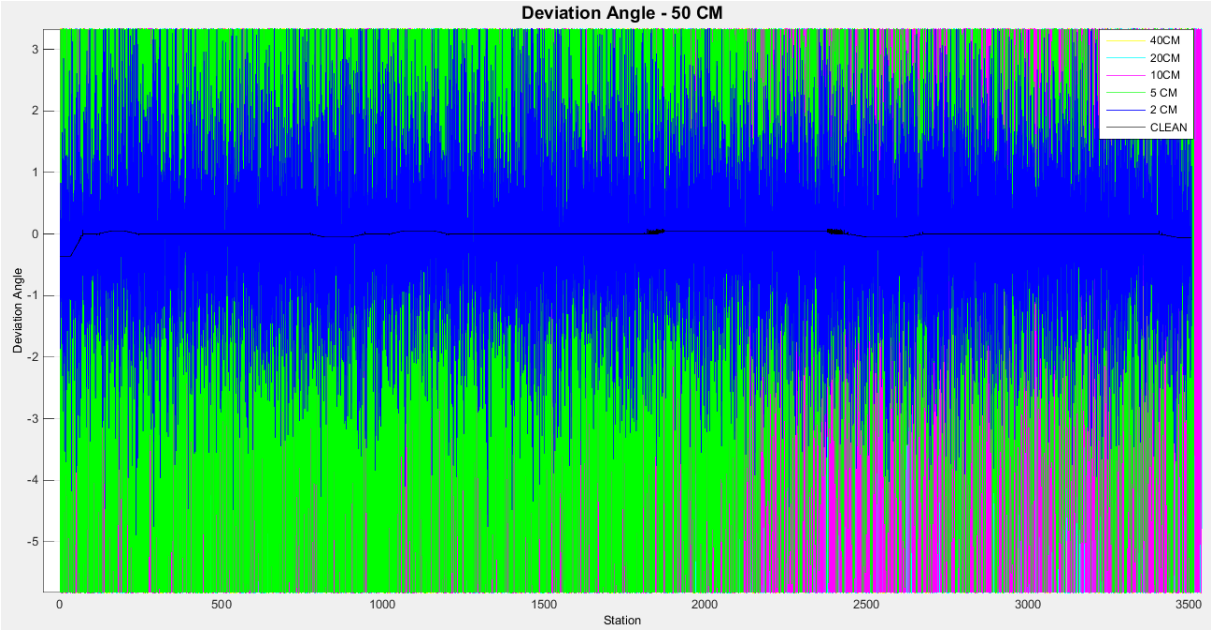


Figure 6.20 - Detailed zoom in deviation angle on design-based track - 50cm

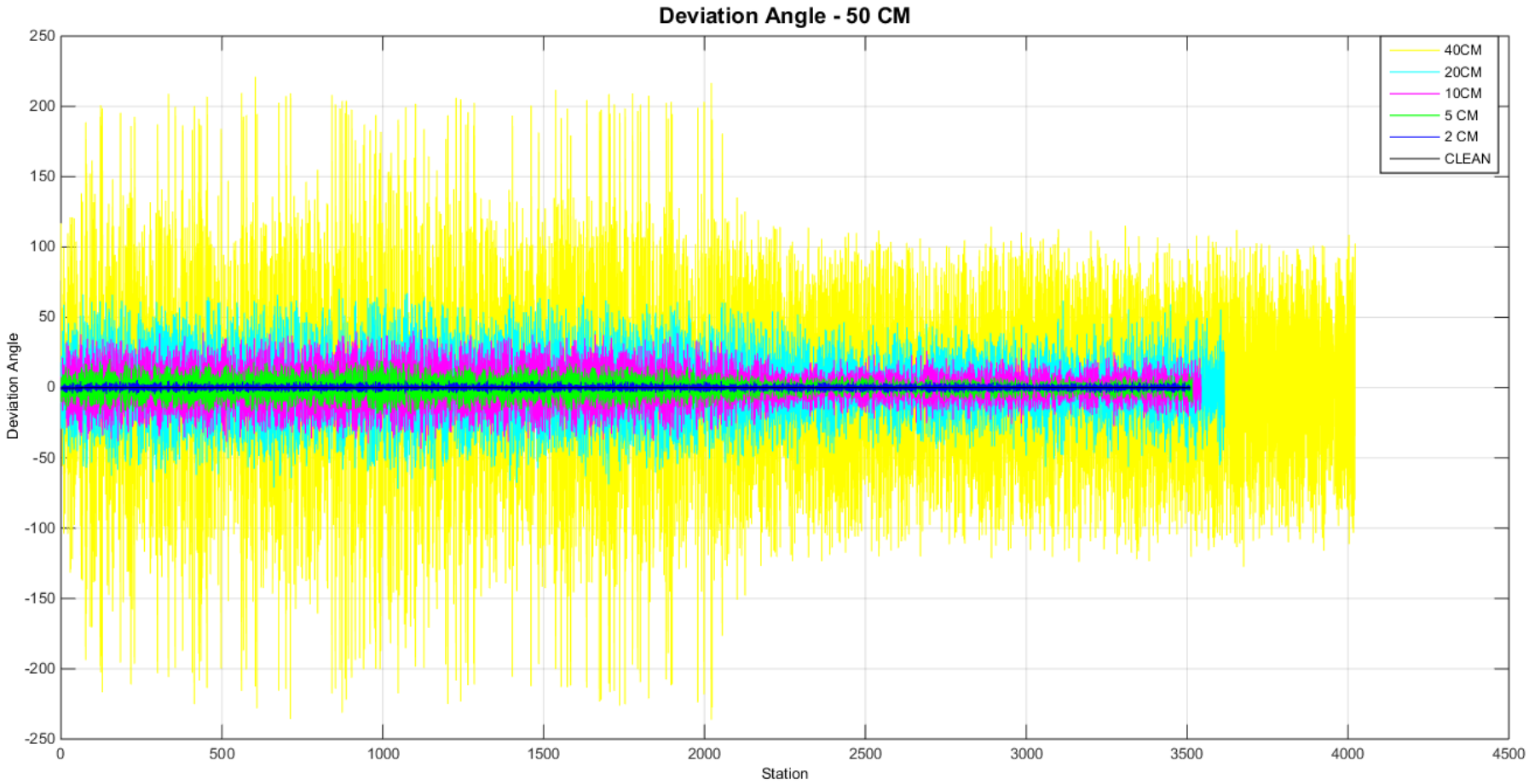


Figure 6.21 - Local deviation angle – design-based tracks – 50cm

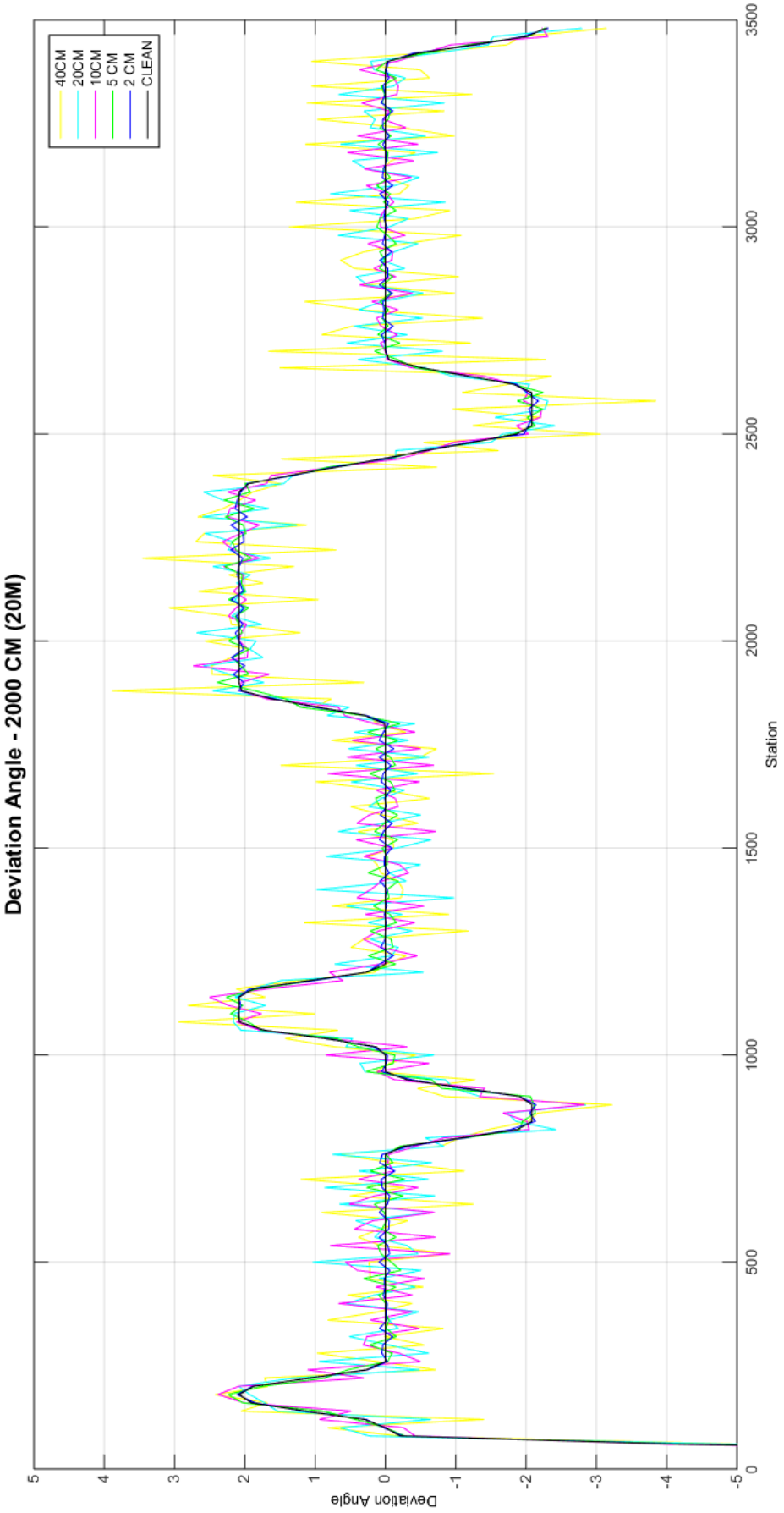


Figure 6.22 – Local deviation angle – design-based tracks – 20 m

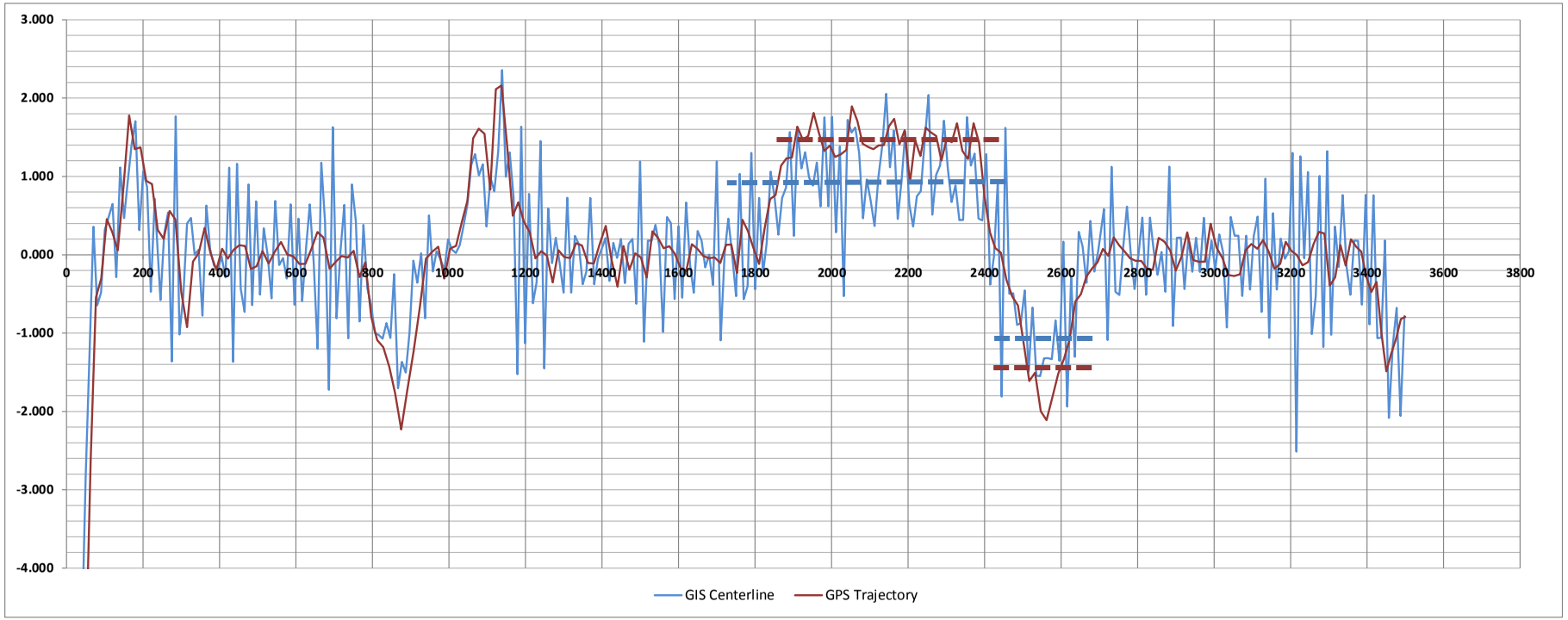


Figure 6.23 – Local deviation angle on survey-based tracks

The real important element that is deductible from these graphs is actually something that is strictly correlated to the potential of being able to automatically detect curves and tangents referring on the deviation angle. Looking carefully at the graph, it is worth noting that, averagely, the deviation angle values on curves seem to be not the same in both GIS-based and GPS-based tracks. Actually, ignoring for a moment the instability of the signal, it is quite clear that the deviation angle values seem to be shifted between the two tracks. In order to better understand this behavior, it has been plotted the deviation angle of tracks with no random error applied, but with different space-based frequency. The results are in Figure 6.25.
In the top part of the picture, there is the deviation angle evaluated on different tracks and it is clearly visible that there is a precise shifting factor between all the tracks. Actually, it is shown that the local deviation angle is a function of the distances between points and the radius of the curve.

In fact, applying a simple threshold method to distinguish between curves and tangents, as sometimes proposed in literature (Li, Chitturi, Bill, & Noyce, 2012), seems not to be valid in general. Actually, it is impossible to set a unique threshold without knowing spacing between points and curve radii (and this last one is specifically one of the final targets, not a datum at all). This error is clear in the bottom of Figure 6.25, where identification is performed fixing a threshold equal to 1.25 degrees (Li, Chitturi, Bill, & Noyce, 2012). Not all curves are identified in the same position due to the deviation angle shifting, and in some specific conditions (typically, with spacing equal to 10 m, 5 m, 2 m, 1 m and 0.5 m) no curves are identified.

There is a specific relationship between the spacing, the radius, and the local deviation angle. This is evincible by looking at Figure 6.24, where a simplified scheme of three successive epochs is shown.

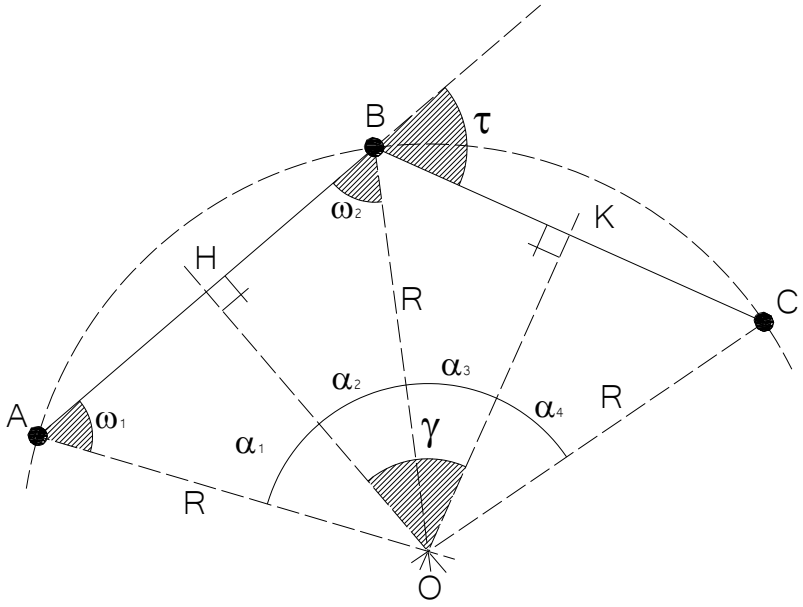


Figure 6.24 – Relation between local deviation angle and central angle

Giving as an assumption that the spacing between points is constant

$$AB = BC \qquad \text{Eq. 6.1}$$

and considering the geometry of the problem, it is possible to state that

$$\overrightarrow{AB} \perp \overrightarrow{OH} \quad and \quad \overrightarrow{BC} \perp \overrightarrow{OK} \qquad \text{Eq. 6.2}$$

This geometrical consideration allows to the following

$$\tau = \gamma \qquad \text{Eq. 6.3}$$

Focusing on one single triangle, like ABO, it is possible to state that is isosceles, since two sides are equal

$$AO = OB = R \qquad \text{Eq. 6.4}$$

allowing to say also

$$\omega_1 = \omega_2 \qquad \alpha_1 = \alpha_2 \qquad \text{Eq. 6.5}$$

Since the two triangles ABO and BCO are equal (equality of three sides), is it finally possible to say

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha \qquad \text{Eq. 6.6}$$

and finally getting the final result

$$\tau = \gamma = 2\alpha \qquad \text{Eq. 6.7}$$

The exact value of τ is

$$AB = 2(R \cdot \sin\alpha)$$
$$\tau = 2\alpha = 2\arcsin\left(\frac{AB}{2R}\right) \qquad \text{Eq. 6.8}$$

It is fundamental to highlight the Eq. 6.1, where it is specifically requested that the spacing between points **must be constant**. However, thinking about mobile mapping surveys and the ways with which they are performed, it is immediate to understand that this hypothesis cannot be respected at all. All common instruments work with a time-based frequency, typically expressed in Hertz, and this is what happened also in our case study, where a specific sampling frequency has been set up before starting the survey (50 Hz on IMU and 1 Hz on GPS). Technically, it could be possible to translate the time-based frequency into a space-based frequency knowing the vehicle speed. This could not be a problem, since generally it is possible to record vehicle speed epoch by epoch, so the conversion is theoretically (and practically) possible.

Nevertheless, obviously, it is not possible to state that the vehicle speed is constant all along the survey; so what about the signal?
Since the speed is not constant, the space-based frequency is not constant too. This means that, epoch by epoch, the theoretical (and exact) deviation angle is not constant too.
That is the explanation of what is shown in Figure 6.23: a random noise with general trends, both of them are due to a continuous balancing between the positioning accuracy, the accuracy of image analysis and the speed variation of the vehicle.

In Figure 6.26 is plotted the Eq. 6.8 with various radii and spacing, and it's clear how much variable the deviation angle is with respect to these two parameters. Please note that both axis (X and Y) are in a logarithmic scale. In the picture are highlighted some specific radii (550 m, 500 m, 80 m) and each dot represent a typical deviation angle value that has to be found in the case study.

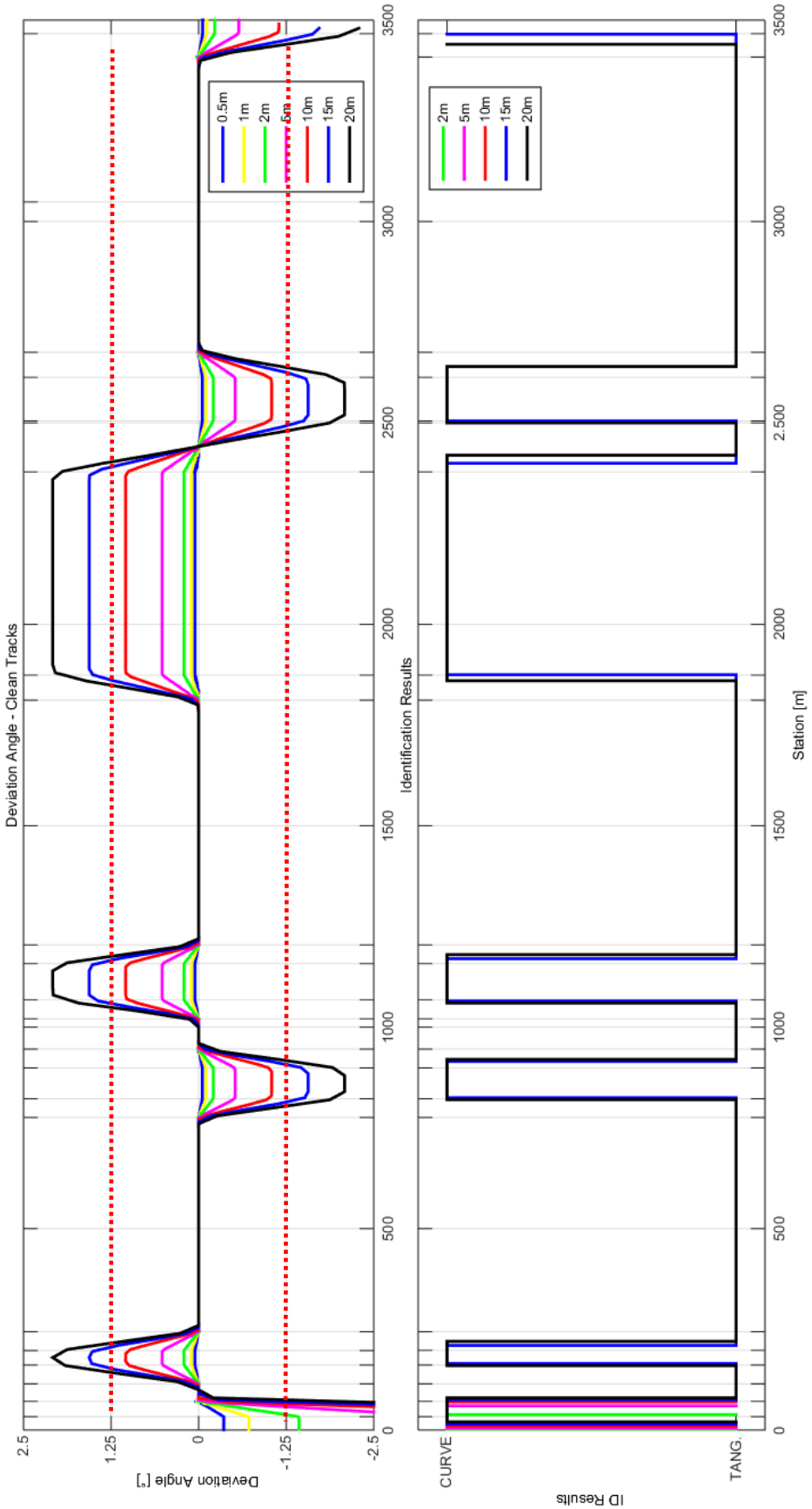


Figure 6.25 – Local deviation angle on design.based tracks with no random error and various spacing

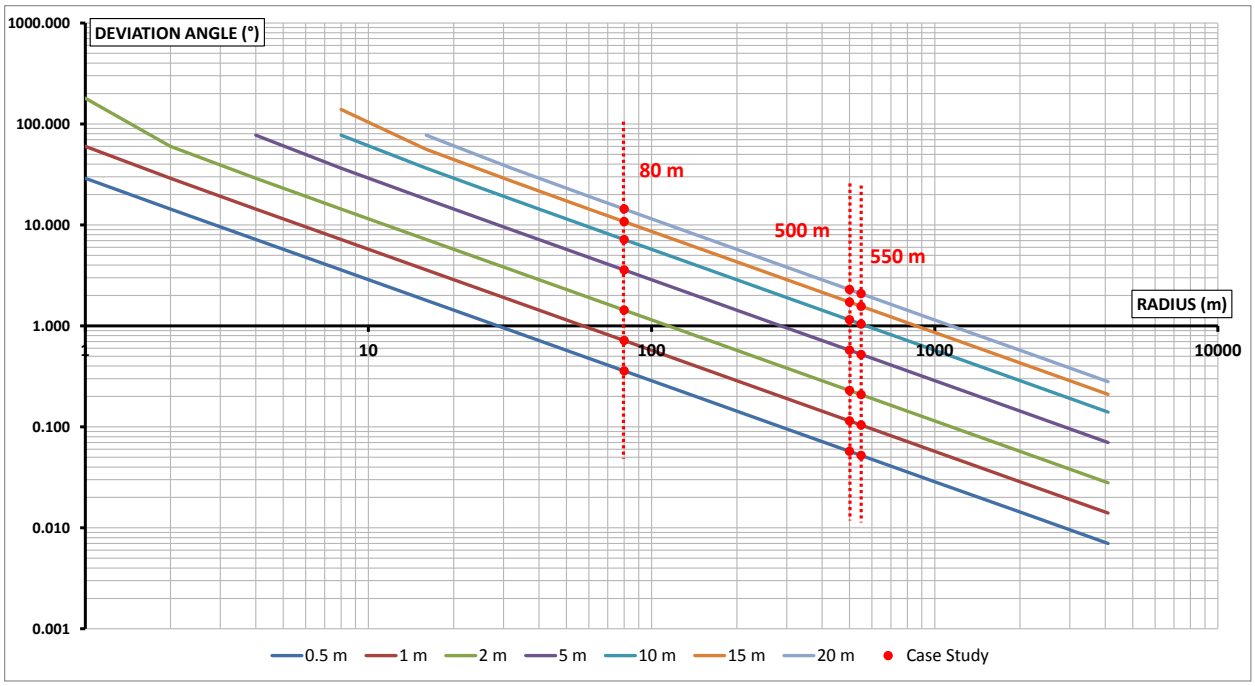


Figure 6.26 - Local deviation angle trends respect to radii and spacing

6.2 GEOMETRY CHARACTERIZATION

6.2.1 Tangents Fitting

The first element that has been fitted is the tangents angular coefficient (m). in fact, tangents are the easiest and fastest to be estimated, since except from singular points, any inaccuracy on point positioning is quite well absorbed by the fitting algorithm itself and the result is generally good. In the following pages, some graphs show the fitting results about main tangents. The case study was composed by seven tangents, seven curves, and twleve transitions curves. In order to not waste too many pages with similar graphs, only few of these fitting results are shown in this paragraph, and a brief comment is added. For the complete list of fitting results, please refer to the attachment "*Fitting results*"

Each single group of graphs (typically histograms) are referred to a specific element of the case study, indicated in the figure caption. Each group is composed (generally) by 8 different histograms, showing each single fitting combination: seven of them are referred to design-based tracks and the last one is referred to survey-based track. The single histogram itself is referred to the same space-based frequency (from 0.5 m to 20 m), containing all different maximum random errors (from the "clean" track – the original one – to a maximum error equal to 40 cm). Finally, for each single combination is available two or three fitting: least squares (LS), Huber robust fitting (HB) and Landau (LD). Landau is not available for tangents and clothoids because it is specifically aimed to fit circles. In each histogram, a dashed line shows the theoretical expected value.
The last histogram, referred to survey-based tracks, contains the fitting for all the surveyed tracks.

Figure 6.27 shows all the fitting results about the first long tangent of the track. The fitting quality is particularly good and in all cases, the angular coefficient of the correct tangent is quite well fitted. Generally, Huber and Least Squares techniques stand on the same result, and generally, the result quality is not reduced significantly from the point positioning accuracy. From the “clean” track to 40 cm, the fitting keeps generally the same result, with no sensitive variations. Fitting errors are distributed equally above or under the correct value and looking at the y-axis variation, the azimuth is well fitted in all situations. The fitting quality generally decreases while the number of points reduces, but this generates a variation of the angular coefficient with the magnitude of the 10^{-3} power (or, in percent, a variation equal to 0.1% of the correct value).

Survey-based tracks are generally less precise than design-based tracks and lead to variations equal to 1% of the final fitted value. Given as accepted that an error equal to 1% of the expected value is quite irrelevant, it is worth noting that all the survey-based tracks overestimate the angular coefficient. Again, no sensitive differences are available between least squares and Huber technique. Similar results are obtained on the second tangent in terms of length (#15), which results are shown in Figure 6.28.
Figure 6.29 shows fitting results of the shortest tangent of the case study truck. It is an extremely short tangent between two clothoids and actually, in almost all cases there are not enough points to be fitted with any method available. Only 50 cm and 1 m spaced tracks have been fitted and it is obvious that the fitting quality is definitely poor and insignificant.

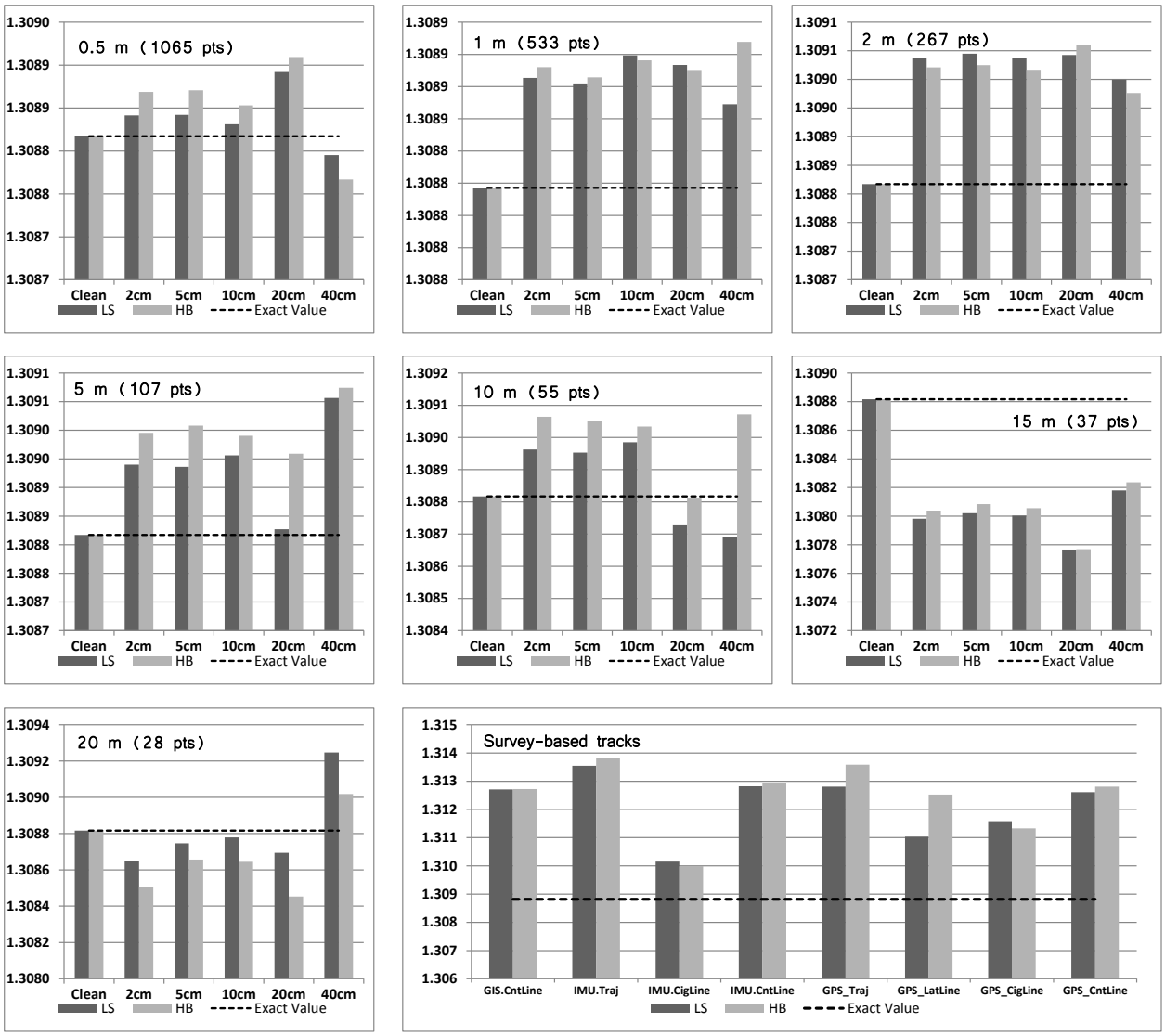


Figure 6.27 - Tangent #7: LS & HB fitting results (m=1.3088)

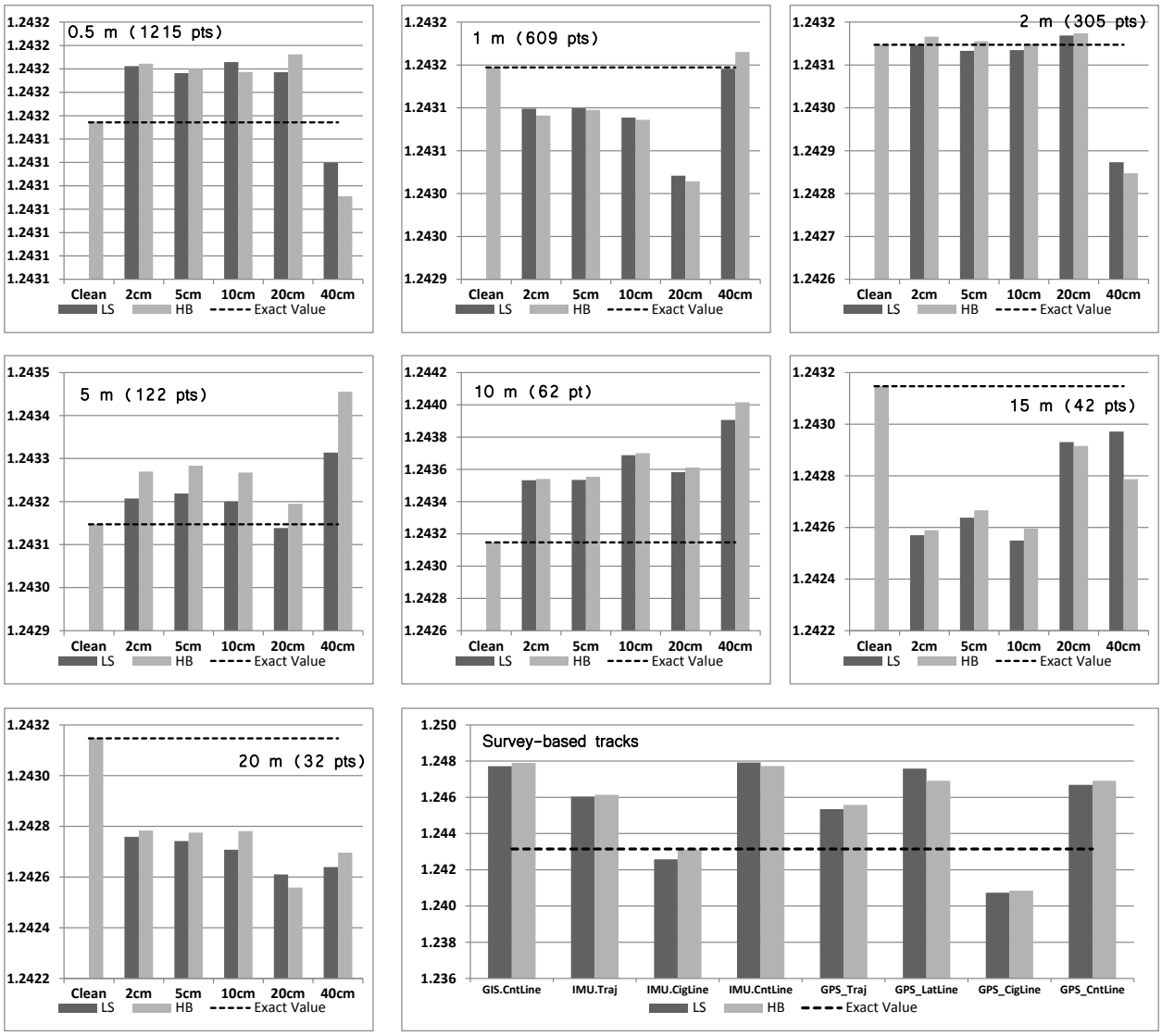


Figure 6.28 - Tangent #15: LS & HB fitting results (m= 1.2431)

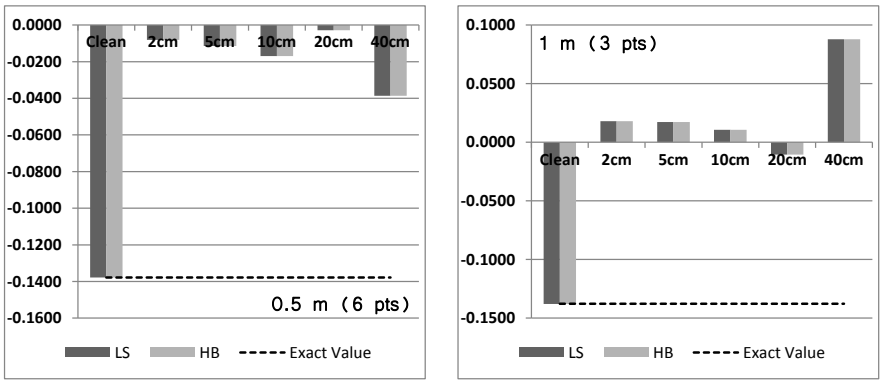


Figure 6.29 - Tangent #19: LS & HB fitting results (m= -0.1378)

In Figure 6.30 there is the last tangent of the track (#23), which is quite long too and well defined. In addition, in this case the fitting is very well performed and errors represent less than the 0.1% of the expected value. In the case of 10 meter based tracks, there is a visible difference between the Least Squares fitting and the Huber one: this effect is attributable to errors randomly applied along perpendicular directions with respect to the tangent orientation. In fact, if these random errors are casually applied along these directions, errors affected points represent a sort of outliers and, consequently, Huber techniques can detect these values and take some adjustments, giving a different fitting. On survey-based tracks, the fitting is quite satisfactory and the errors are below the 0.1% of the desired value.

It is possible to state that tangents fitting do not represent a problem in terms of results quality and fitting algorithm effectiveness.

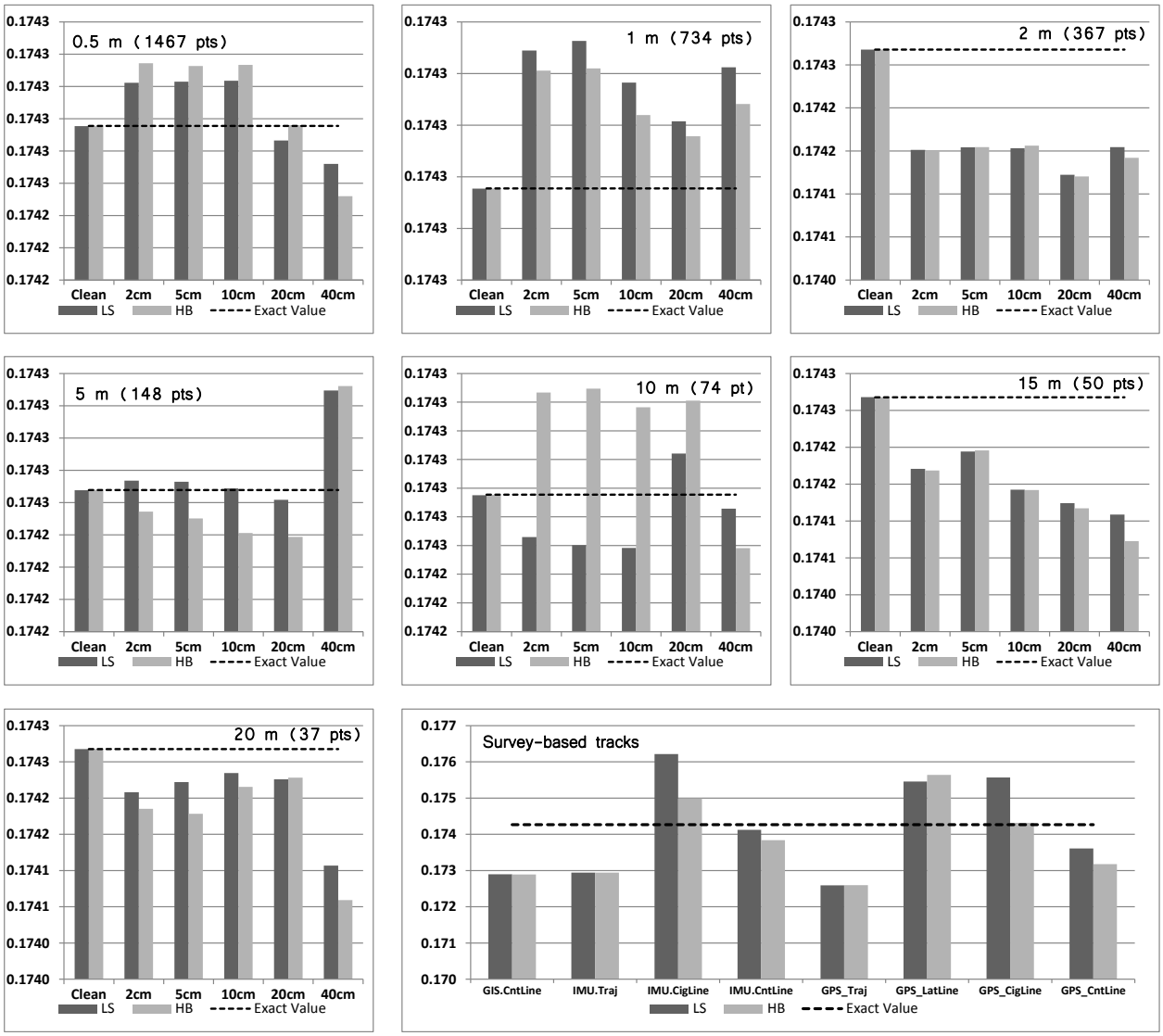


Figure 6.30 - Tangent #23: LS & HB fitting results (m= 0.1743)

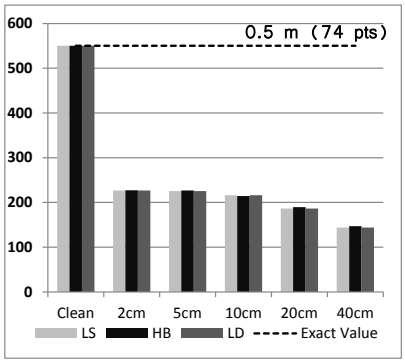
6.2.2 Curves Fitting

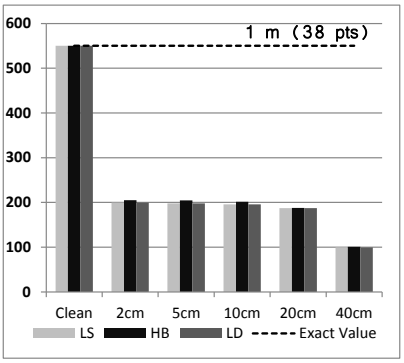
Curves are definitely the most difficult and challenging element to be fitted. No problems occur in tangents fitting, due to the geometry of the problem and because fewer points are needed to obtain a good redundancy. Curves instead need more points and generally are sensitive to small variations: small errors on points can heavily generate some errors on center positioning and radius.

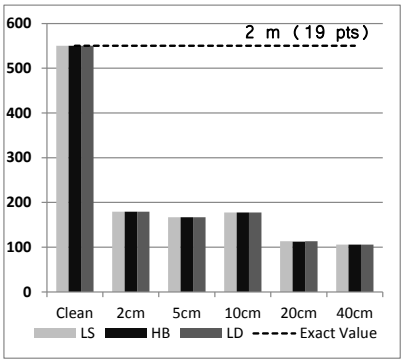
Similarly to tangents, in the following pages a selection of fitting results are shown, and in particular fitting graphs related to curves not completely available for the fitting are omitted. The first and the last circular curve (#1 and #25) were not completely available from the survey, since the start of the survey and the end were along the circular element. Given that loss of points, they are not supposed to be the best fitted among all. Anyway, the complete set of fitting graph is always available in attachments.

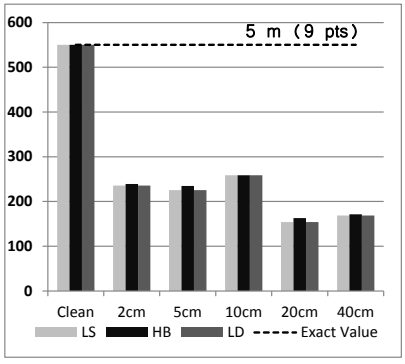
Figure 6.31 shows the first complete circular curve of the track (#5). This curve is included between after a short tangent, connected with a transition curve; this curve is quite short, with a central angle equal to 3.82°. Looking in detail inside the picture, it is possible to state that the clean track is always perfectly fitted, without any correlation with the number of points or with the technique adopted. When a random error is applied, the fitted radius is significantly smaller than the expected one, and this reduction from the exact value tends to increase when the error is more and more big. There are no visible differences between least squares, Huber, and Landau fittings, while there is a clear common tendency to obtain a smaller fitted radius with the increase of the maximum error. This general tendency is no more “respected” when boundaries fitting conditions are changed, like when there is a big reduction of the number of points and the redundancy of the system starts to be small.

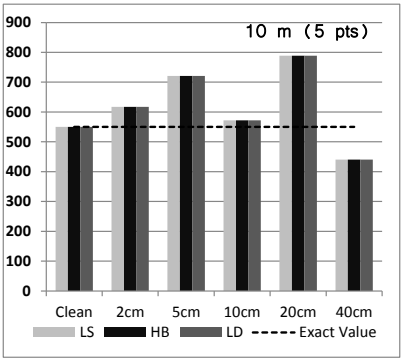
In order to fit a circular curve, it is necessary to estimate four different parameters, each one is included inside the circumference equation. In order to perform a fitting estimate with a solid redundancy and a real possibility to equalize any possible inaccuracy it is necessary to have a set of points definitely larger than the minimum needed. With four parameters to fit, five points are the minimum number requested to avoid a single-determined solution. This means that when this threshold is really close to the real point number available, the fitting result is not only dependent from the fitting algorithm or from the point accuracy, but it also depends from the ability of the fitting process to compensate errors and lead to the real best fit solution. In Figure 6.31 this phenomenon is clear with 10 m spacing. In this condition, only 5 points compose the curve #5 and so the general tendency to obtain a smaller radius than the expected one is not “respected”. Furthermore, fitting was not permitted with bigger space frequencies (15 m and 20 m) due to the lack of the minimum number of points needed to perform a regular fitting.











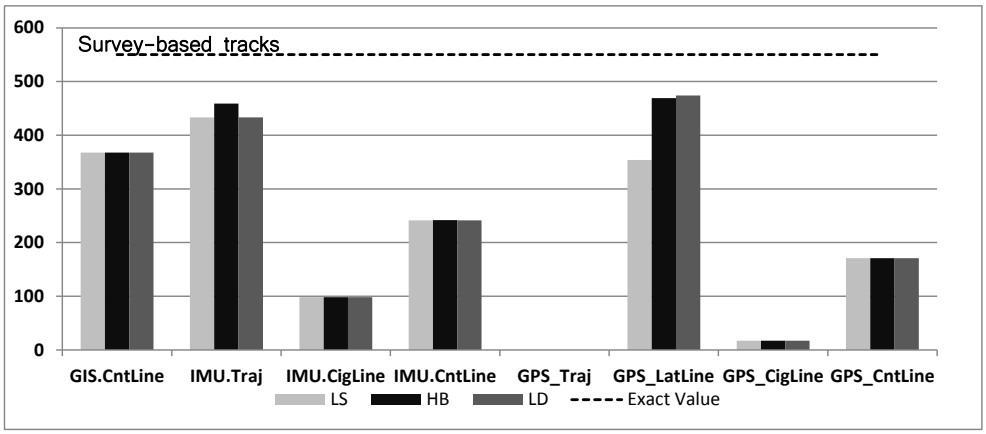


Figure 6.31 - Circular curve #5: LS, HB & LD fitting results (R= 550m)

To confirm this hypothesis, it is possible to look at other curves fitted, like ones presented in Figure 6.32, Figure 6.33 and Figure 6.34. These graphs are related to curves #9, #13, and #17, with respectively a central angle equal to 7.98°, 9.44°, and 52.55°.

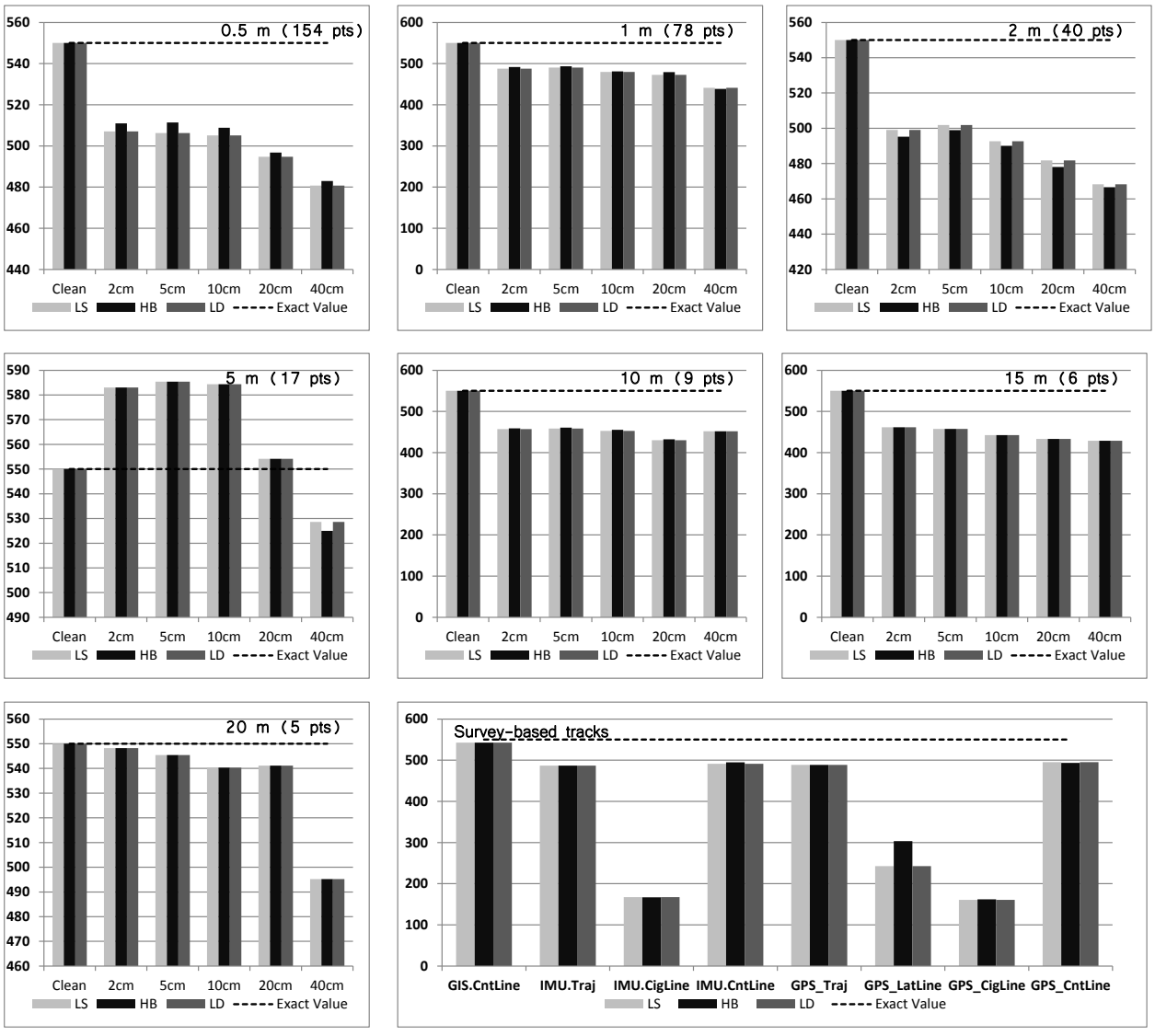


Figure 6.32 - Circular curve #9: LS, HB & LD fitting results (R= 550m)

It is possible to state that there is a general tendency that is respected from all the case study curves shown in these pages. Except from the clean tracks, which are always correctly fitted, all the other combinations give back a fitted radius that is smaller than the expected value. Furthermore, the effect of the maximum random error applied to each track is not definitely relevant with respect to the average reduction of the fitted radius. Point positioning quality produces a fitting error that is not comparable with the average overall fitting error: actually, it seems that the quality of the fitting is not leaded by the point positioning or from the point number, but is dependent from another parameter.

To discover the actual parameters that play a consistent role in the fitting precision it is necessary to look carefully at fitting results obtained on the curve #17. Actually, there is no substantially difference between the other graphs, but it is necessary to focus on y-axis scale to understand the real difference. Even if there are some differences between all the parameters combinations (spacing, maximum random error, and fitting technique) it is clear that the fitting results are impressively sharp and close to the theoretical value. Even in worst case, with a big spacing and a big error, the radius estimate is definitely good and the difference between the expected value and the fitted one are smaller than 15 cm (that is 3% of the correct value). All the fittings, in all conditions and regardless of positioning accuracy, are extremely precise and easily reach to individuate correctly the radius of the circular curve. Therefore, the question is why this specific curve is so well defined while others are not defined as well: this difference is not ascribable to positioning error, fitting method or space frequency.

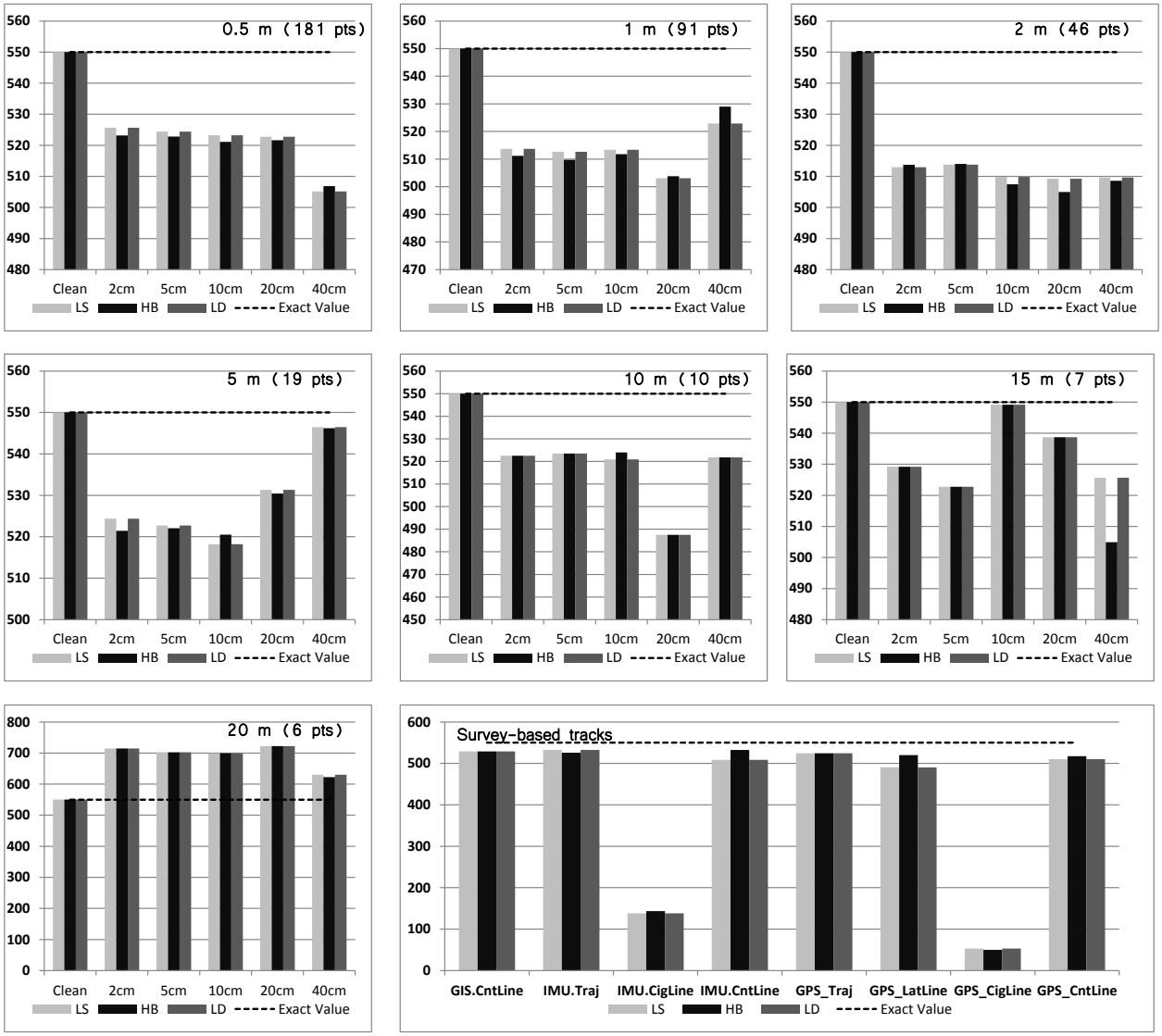


Figure 6.33 - Circular curve #13: LS, HB & LD fitting results (R= 550m)

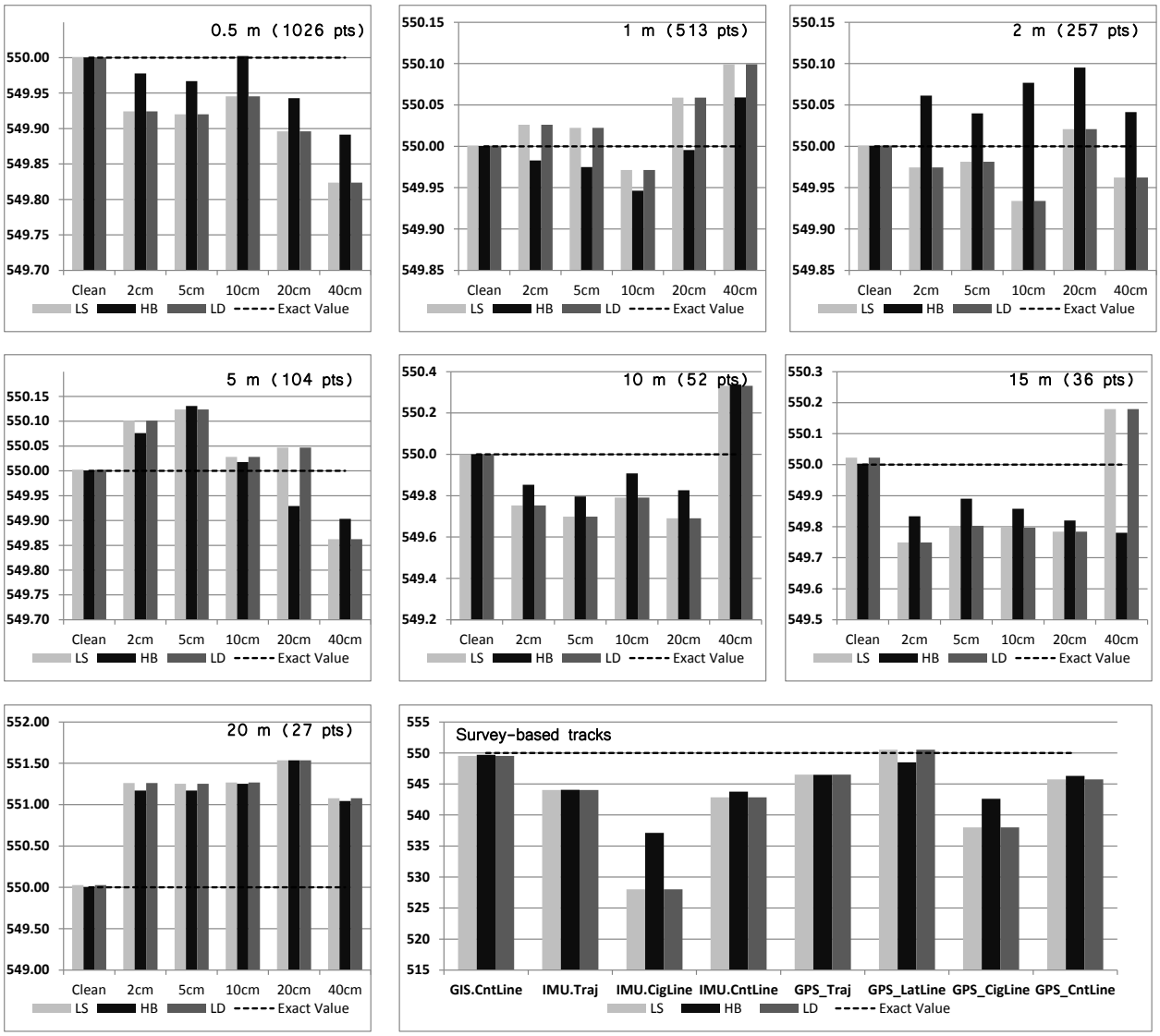


Figure 6.34 - Circular curve #17: LS, HB & LD fitting results (R= 550m)

The final answer is available in Figure 6.35 (survey-based tracks), Figure 6.36 and Figure 6.37 (design-based tracks), where are shown the fitted radius versus the central angle of each single curve. It is clearly recognizable a trend that links the fitting quality with the central angle of the curve. The bigger is the central angle, the better is the fitting, with poor changes between random errors applied and space frequency.

This result demonstrates that the wideness of the central angle of the curve is a predominant factor among all the parameters that have been tested in this work. With accurate positioning and even changing the space frequency between points, the main parameters that has to be taken into account is the length of the curve and its radius, or in other words the central angle.

To find some guidelines to manage this problem and to propose a solution to practitioners and stakeholders of high-speed road survey techniques, a study about this phenomenon has been performed, explained in 6.3.

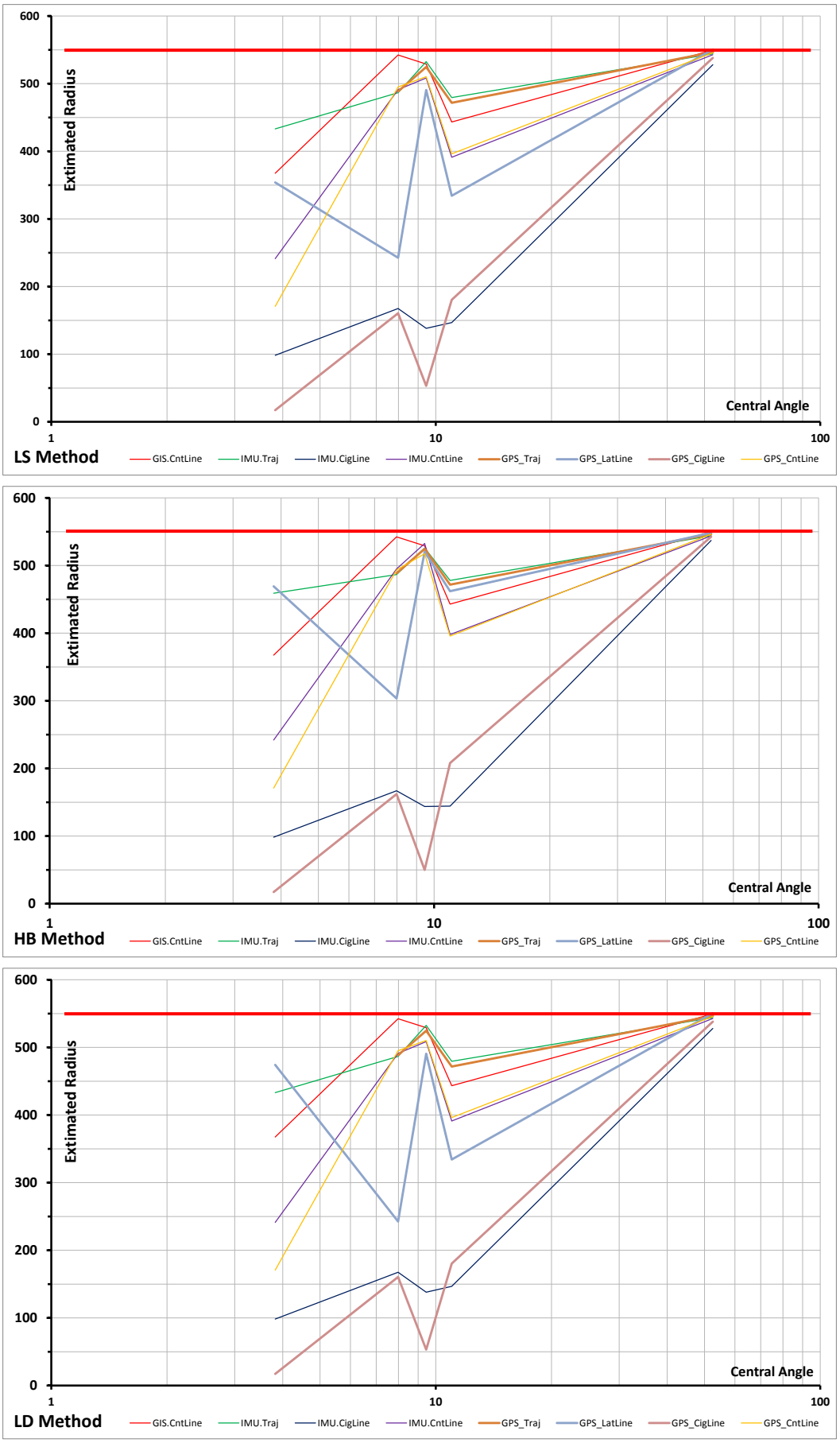


Figure 6.35 - Survey-based fitting radius versus central angle

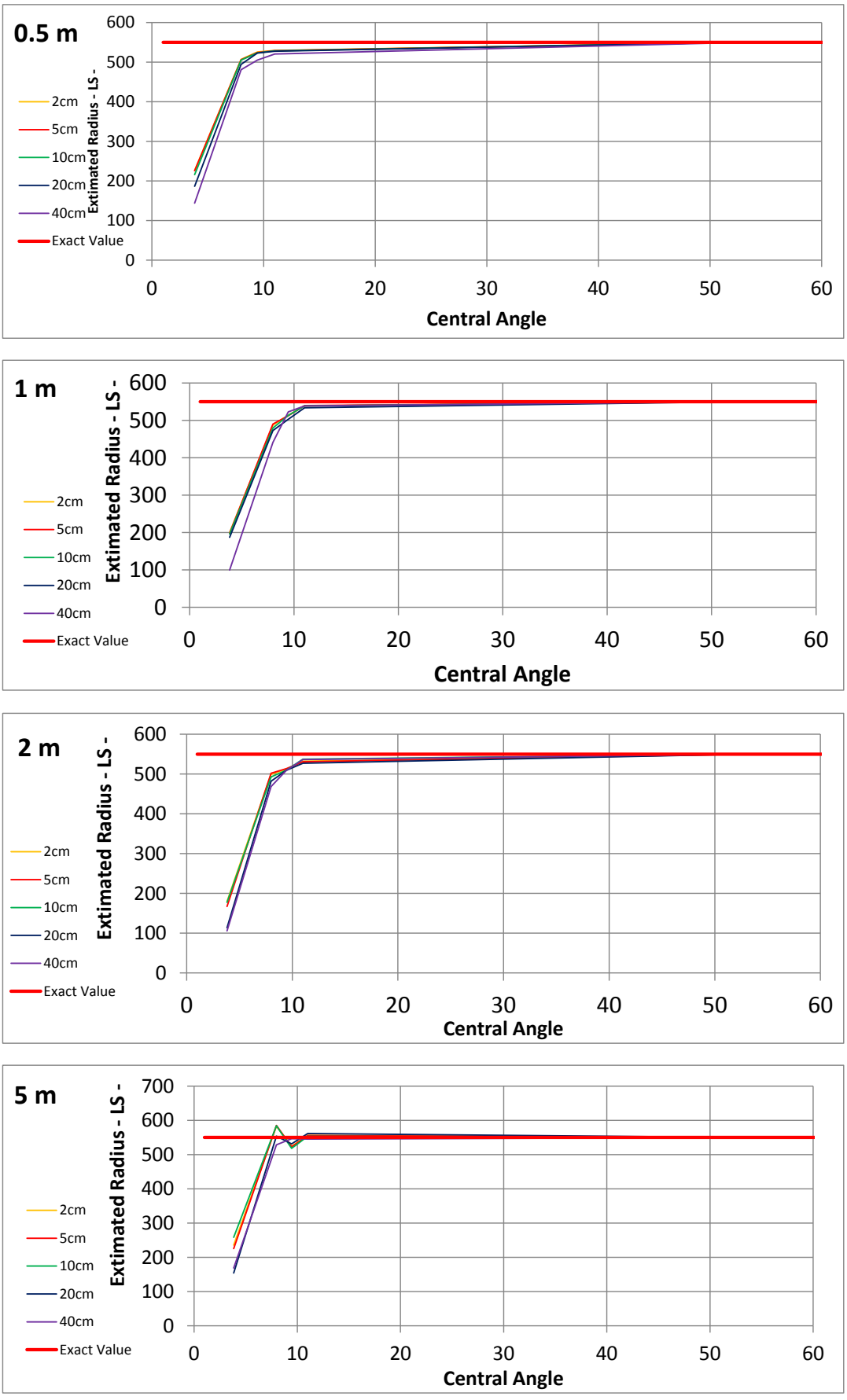


Figure 6.36 - Design-based fitting radius versus central angle (1)

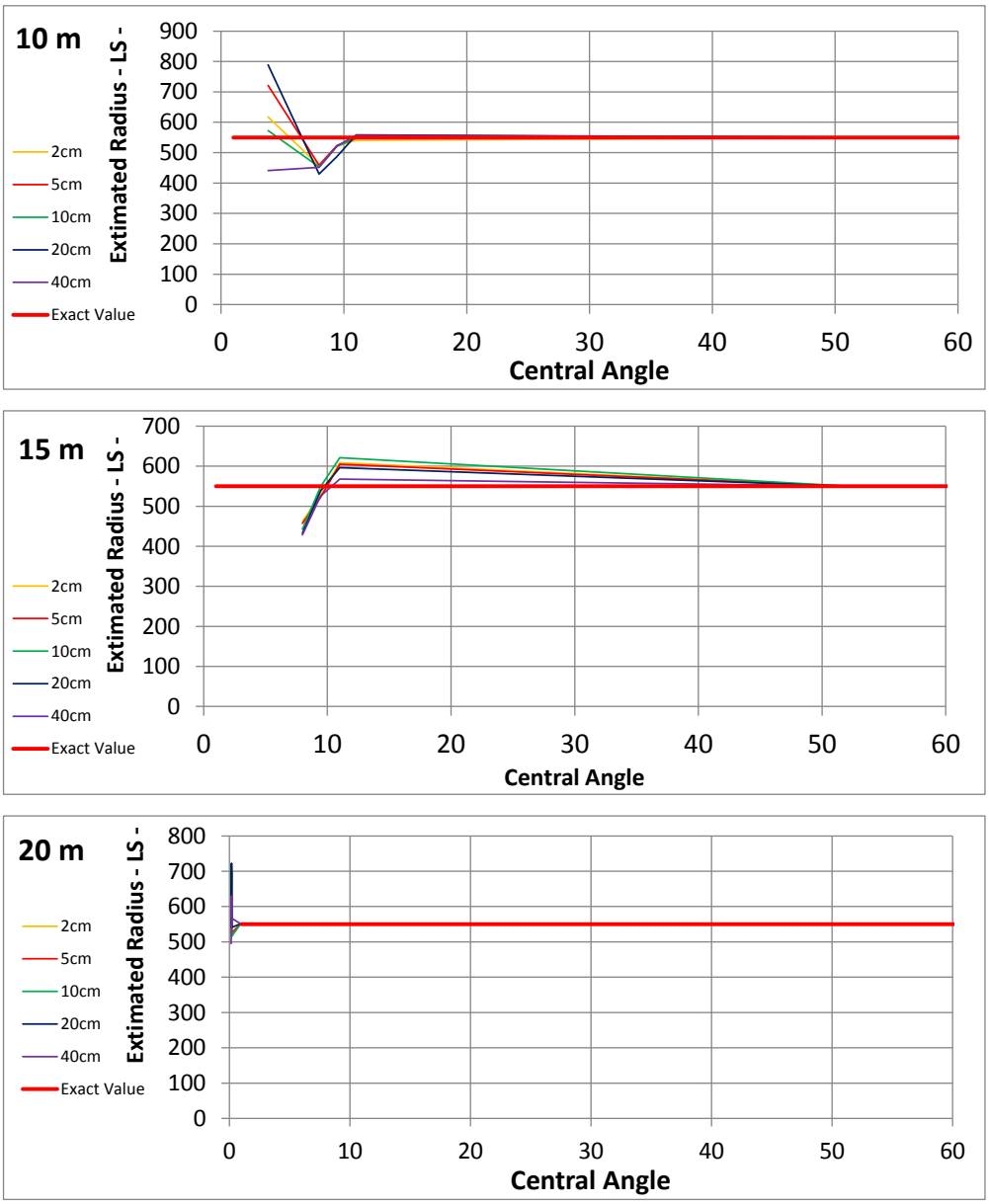


Figure 6.37 - Design-based fitting radius versus central angle (2)

6.2.3 Clothoids fitting

Nothing relevant has to be illustrated about clothoids, since these geometrical elements are purely dependent on the quality of adjacent element of the track: if the tangent and the curve (or both curves in a flex or continuity clothoid) are properly defined and fitted, the clothoid will be properly defined as well. The only important thing that has to be highlighted about clothoids is that it is possible to encounter two different errors while fitting them:

Adjacent elements not properly defined: This error can occur when any of the two elements close to the clothoid has not been properly defined or fitted. This situation is generally due to a not-present fitting on the nearest curve (maybe due to a lack of points in the curve) because the tangent fitting is generally well performed and it gives back good results;

ΔR is negative: This error is generally due to a lack in the fitting of the curve, when the radius has been badly evaluated and it ends up with no displacement between the curve and the tangent. In fact, as well known, the transition curve can only be inserted when there is a pre-determined distance between the tangent and the curve, and if the radius of the curve is not properly fitted, it can happen that the curve intersects the tangent. This condition is obviously impossible and the clothoid cannot be evaluated.

Summarizing, clothoids follow the quality of the fitting of adjacent elements. In Figure 6.38 an example of fitting results, while the complete list is available in the attachment "*Fitting results*".

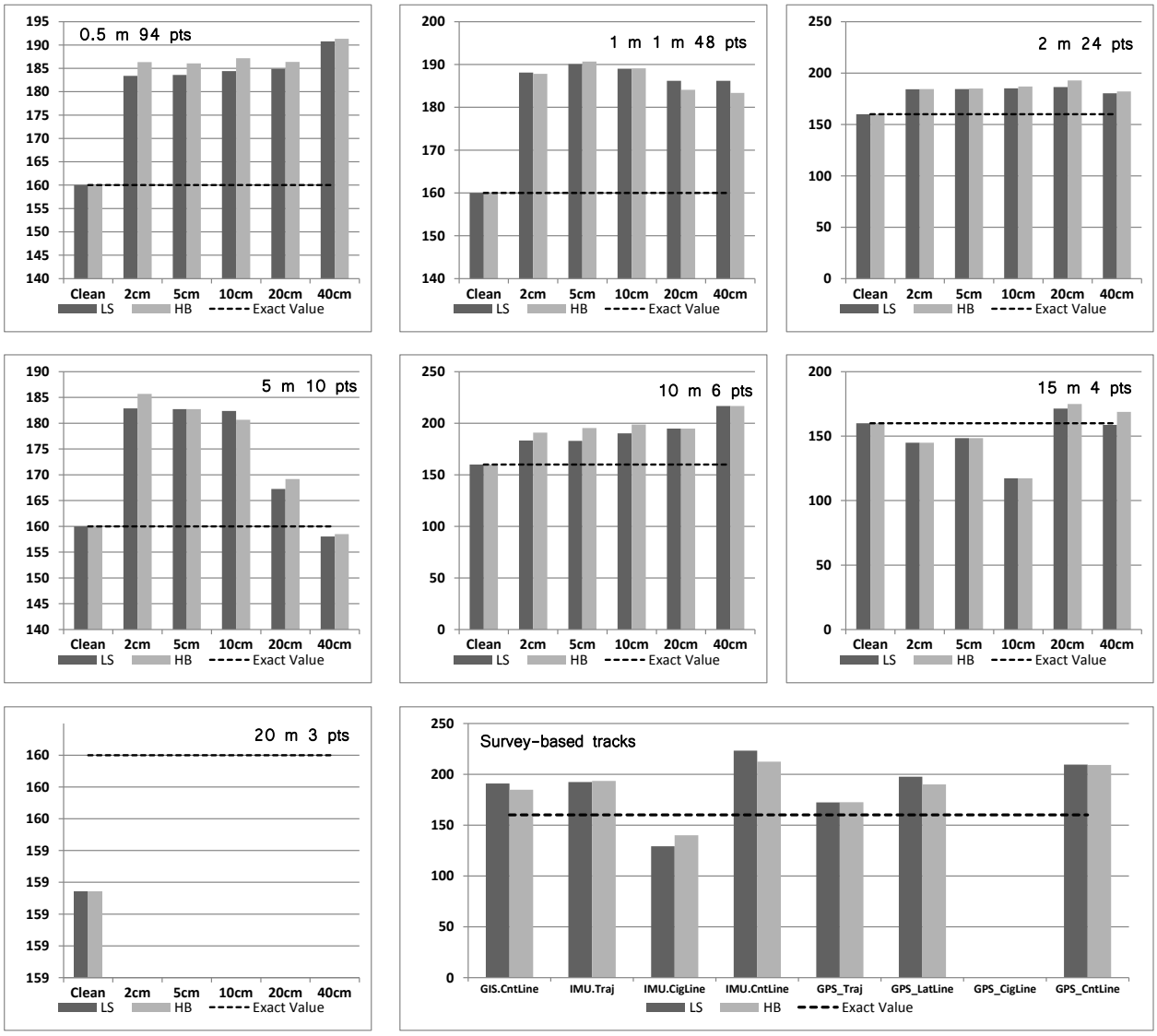


Figure 6.38 - Clothoids fitting example

6.3 BAD CONDITIONING SIMULATION

In order to study the effect of the central angle into the radius fitting, it has been performed a dedicated study with the development of an additional MatLab© code. The logical purpose of this code was to simulate a circle arc fitting while the central angle slowly increase from a minimum to a maximum value, evaluating the trend in the fitting results. The processes of discretization and fitting of data points have been performed on some reference arcs.

Figure 6.39 reports the flow charts of the algorithms employed for this task.

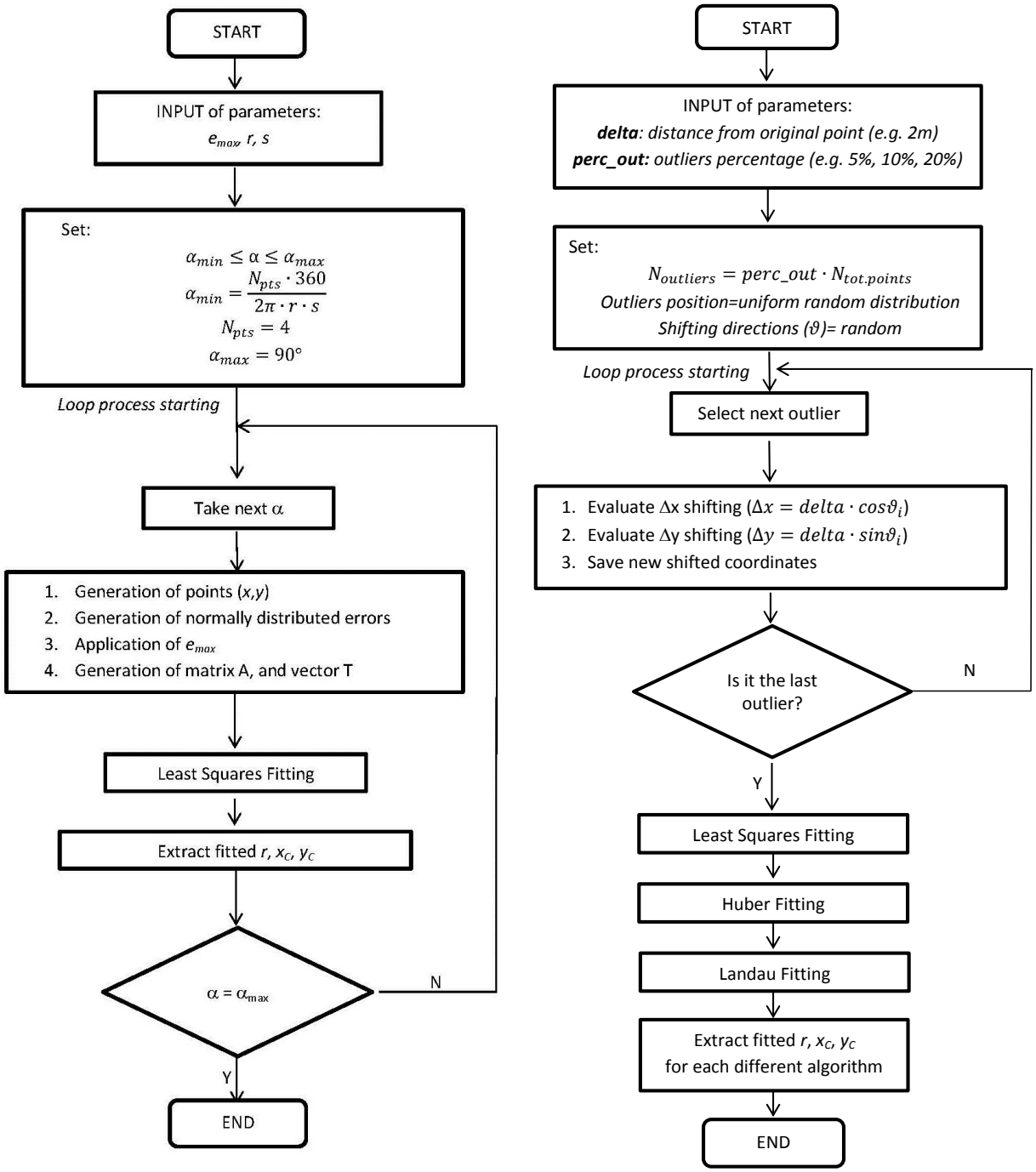


Figure 6.39 – Flow chart of bad conditioning algorithms (a – random normally distributed errors, b – outliers)

The dataset of points generated in the simulation is extracted along circular arcs, as indicated in Figure 6.40, after selecting the parameters such as the spacing between points (**s**), which represents the average distance measured

along the curve from points extracted from the arc, the radius **r** and the central angle (α). Points are generated varying α from a minimum (α_{min}) to a maximum value (α_{max}). The α_{min} is set 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 are considered in the fitting process (one more than the three researched parameters **a**, **b** and **c**).

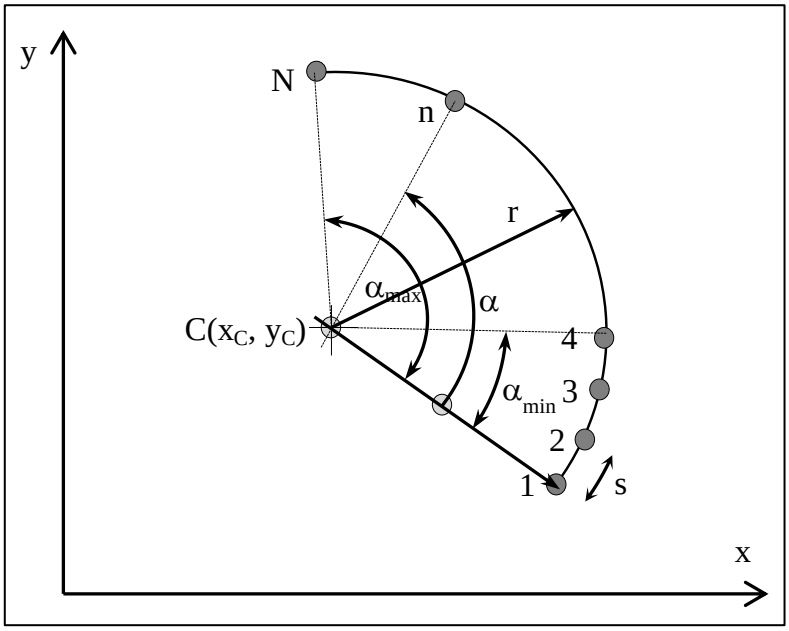


Figure 6.40 - Generation of a sequence of points belonging to a circular arc

Afterwards, each single point of the sequence is moved slightly from its original position by applying two noise terms (Δ_X and Δ_y) through the implementation of a sub-routine in the algorithm. 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 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 [Figure 6.39-b] was then used to determine the new position of those points that were artificially modified to become outliers. The number was selected a priori, the selection and the 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, after 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 maximum angle reaches the maximum value of **90°**.

Every arc was treated with values of e_{max} equal to 2, 5, 10, 20, and 40 cm, and different spacing between points equal to 0.5, 1, 2, 5, 10, 15 and 20 m. In the case of the outliers, the maximum error of 2 m was adopted.

Figure 6.41 contains an example of a circular arc composed of 55 points spaced 5 m apart, and determined with an accuracy of 10 cm, generated according to the procedure described before. Such arc has been fitted with three different curves determined according to the LS, Huber and Landau methods..

No disparities in the results are evident from an analysis of the graph, while data reported in Table 6.1 evidence the differences in the estimation of a, b and c parameters, the radius and the coordinates of the center of curvature. Of the three solutions, the LS and Landau are practically identical, while the Huber one is slightly different and closer to the original circle. The graph in Figure 6.42 contains the fitting of the same data points that were previously modified including 20% of outliers, while Table 6.1 also reports the complete set of results for a presence of outliers equal to 5, 10 and 20%.

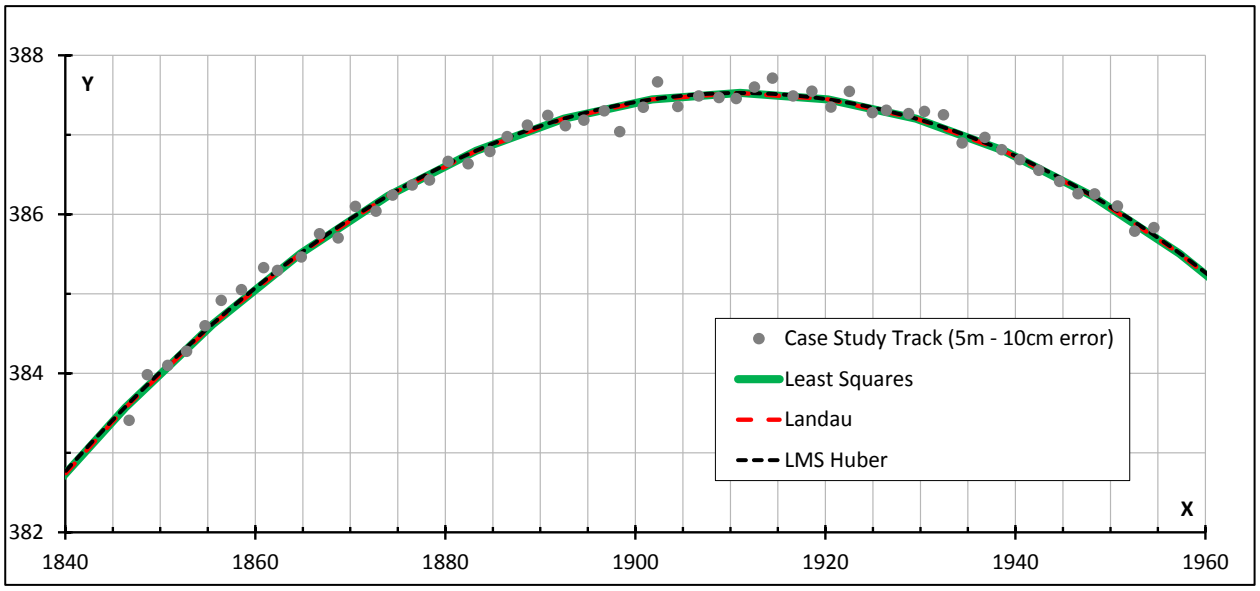


Figure 6.41 – LS, Huber & Landau Fitting (5 m spaced points / 10 cm error)
All points are belonging to a curve of 108 m in length and an original radius of 550 m

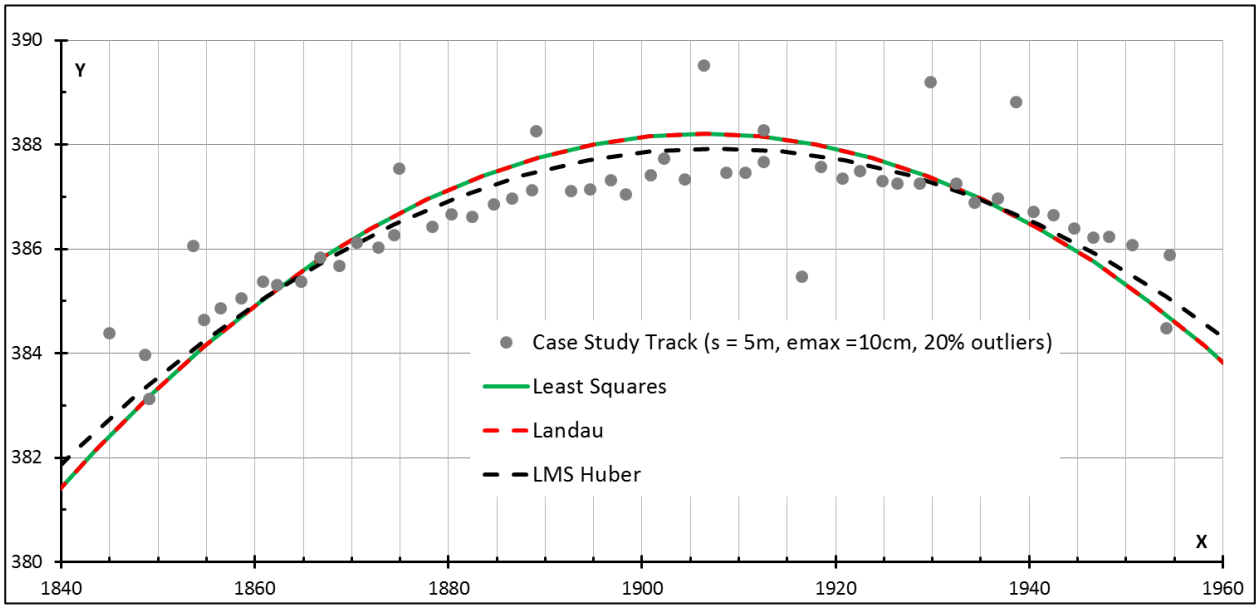


Figure 6.42 – LS, Huber & Landau Fitting (5 m spaced points / 10 cm error / 20% outliers)
All points are belonging to a curve of 108 m in length and an original radius of 550 m

Table 6.1 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 $\Delta(r)$ in the table and expressed as a percentage of the reference value of 550 m, become notable when the presence of outliers differs from 0%. It is also worth noting that when the level of disturbance in the data increases the center of the fitted circular arcs tends to move towards the curve, and the radius decreases significantly.

Parameter		a	b	C	x_c	y_c	r	Δ(r)
Measurem. unit		m	m	M	m	m	m	%
Arc properties		**-3,822.478**	**324.929**	**3,376,728.912**	**1,911.239**	**-162.465**	**550.000**	**-**
0% Outl.s	LS	-3,822.034	286.581	3,390,751.724	1,911.017	-143.291	530.816	-3.5
	Landau	-3,822.034	286.581	3,390,751.724	1,911.017	-143.291	530.816	-3.5
	Huber	-3,822.042	289.288	3,389,719.058	1,911.021	-144.644	532.169	-3.2
5% Outl.s	LS	-3,821.101	161.939	3,437,075.513	1,910.550	-80.969	468.704	-14.8
	Landau	-3,821.101	161.939	3,437,075.513	1,910.550	-80.969	468.704	-14.8
	Huber	-3,821.341	241.181	3,406,947.480	1,910.670	-120.590	508.188	-7.6
10% Outl.s	LS	-3,814.704	-23.043	3,496,439.793	1,907.352	11.522	376.411	-31.6
	Landau	-3,814.704	-23.043	3,496,439.792	1,907.352	11.522	376.411	-31.6
	Huber	-3,815.369	-12.777	3,412,094.733	1,907.685	6.388	476.662	-13.3
20% Outl.s	LS	-3,812.933	-116.378	3,529,087.502	1,906.466	58.189	330.019	-40.0
	Landau	-3,812.933	-116.378	3,529,087.502	1,906.466	58.189	330.019	-40.0
	Huber	-3,815.369	-12.777	3,493,727.531	1,907.685	6.388	381.541	-30.6

Table 6.1 – Synthesis of results of fitting for the circular arc of Figure 6.41 and Figure 6.42

Figure 6.43 reports the results obtained for the same target value of 550 m. The x-axis represents the value of the subtended arc (α), while on the y-axis the radius (r) of estimated arc according to the LS method is reported. The graph includes ten different results generated by running the algorithm several times using a 0.5 m of spacing and 40 cm of accuracy. The solutions pertainings to the same running have been linked with the same line.

The results are distributed along an S-shaped curve comprised between the two horizontal asymptotes of 550 m for $\alpha \rightarrow 90°$, and of 0 m for $\alpha \rightarrow 0$. This indicates that when data are affected by casual errors with known levels of accuracy, radii estimated through fitting algorithms are frequently underestimated (in Figure 6.43 only a few points are above the target value of 550 m). 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 6.43, 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 6.44 illustrates the results obtained from the algorithm assuming a radius of 550 m. The first graph (Figure 6.44-A) reports the radius estimates for different values of spacing 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 spacing (between points) does not seem to be of critical importance for the estimation of the radius. It is only when high spacing values (> 5 m) are considered that the number of points available are lower than four 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°.

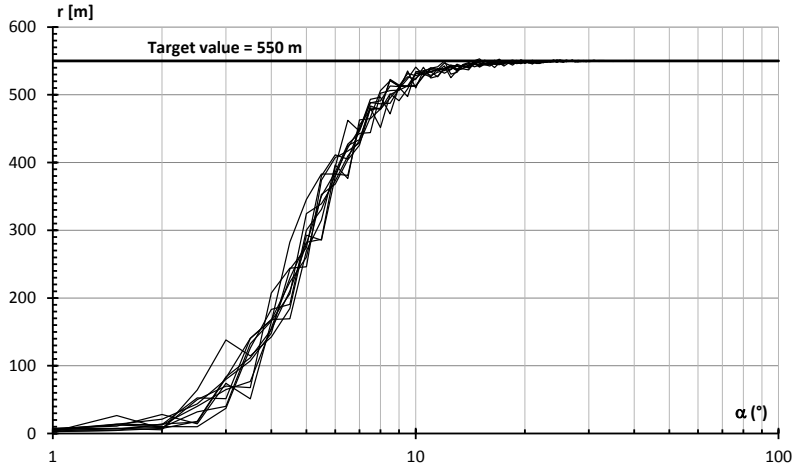


Figure 6.43 – Fitting radii versus the central angle
Spacing between points equal to 0.5 m, and accuracy (e_{max}) equal to 40 cm

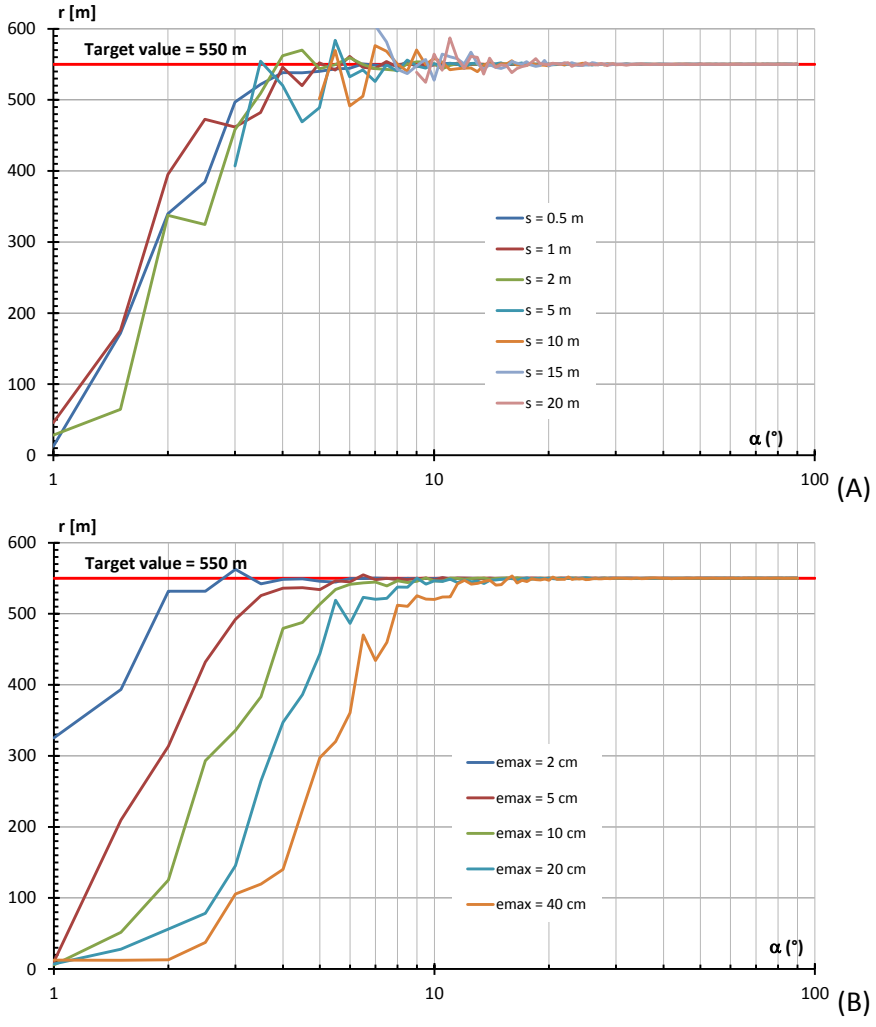


Figure 6.44 – Values of fitting radii as a function of various parameters:
the maximum value of the central angle (α_{max}) for different spacing (A) and accuracy (B)

The curves in the second graph (Figure 6.44-B) have been obtained for a spacing 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 10, 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 spacing of 0.5 m. In the case of small radii, results show that either an increased arc length or greater accuracy in measurements is necessary to attain a satisfactory estimate of the radius. Minor problems are encountered when the radius increases. Where r = 5500 m, no problem of fitting occurs when there is high sampling frequency and accuracy. 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 William, Findley, Zeeger, Sundstrom, & Hummer, 2012).

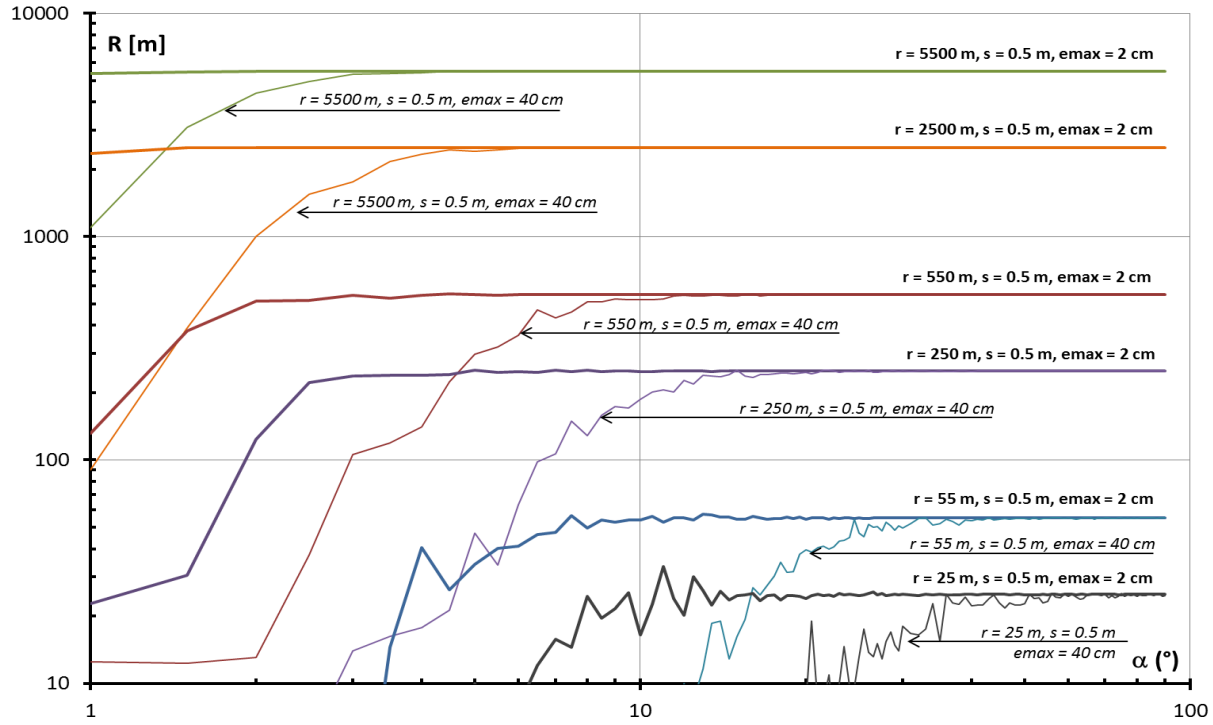


Figure 6.45 - . Values of fitting radii as a function of the radius magnitude (accuracy values - emax - of 2 and 40 cm, spacing of 0.5 m)

Starting from the data in Figure 6.45, in Figure 6.46 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 6.46-A exhibits the regression curves with spacing between data points of 0.5 m, while Figure 6.46-B reports the same for a spacing of 15 m. In both figures, the accuracy value ranges from 2 to 40 cm.

In Figure 6.46-A 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 6.46-B 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.

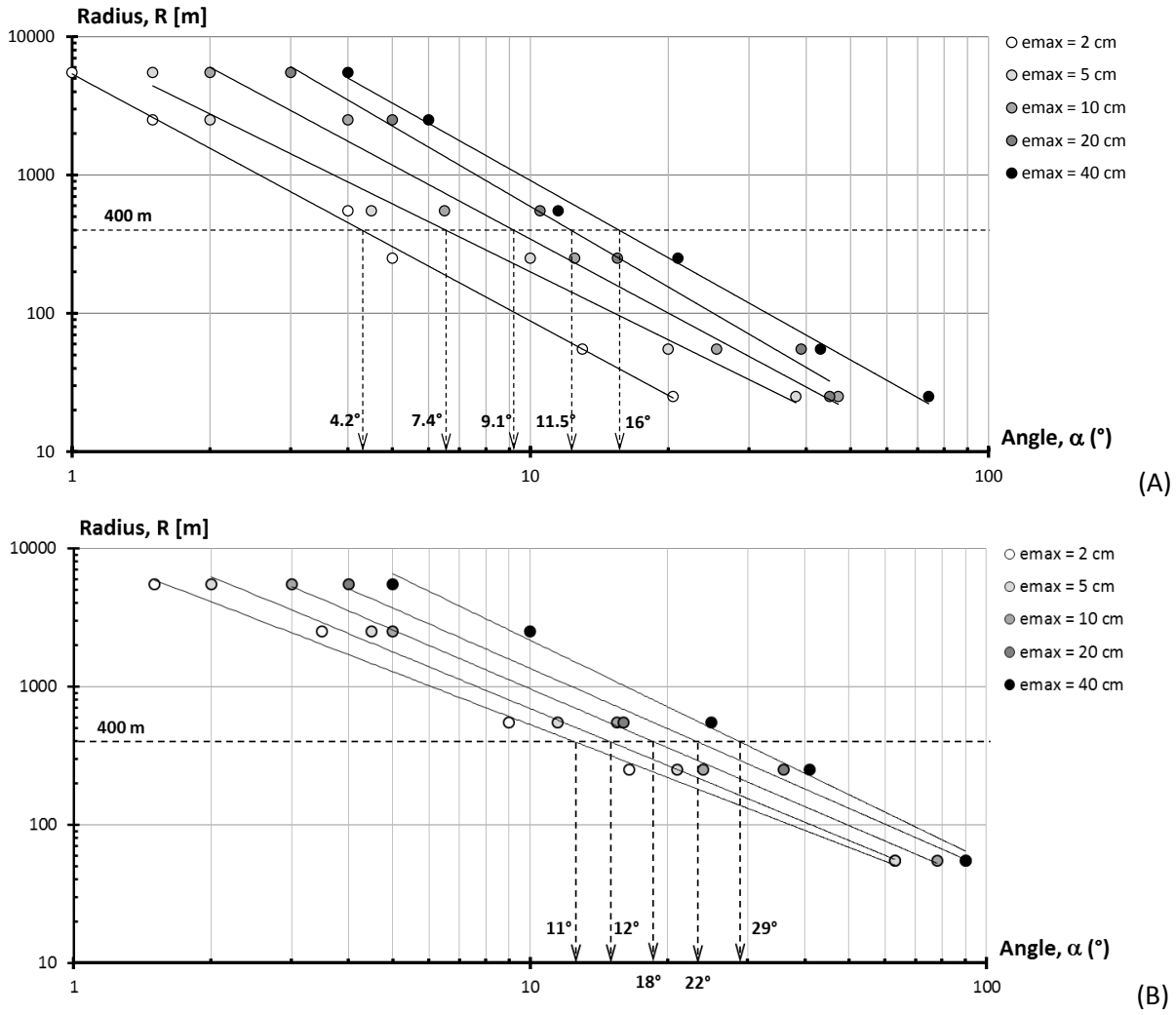


Figure 6.46 - Threshold curves that guarantee a radius estimation error of 2%: a) spacing of 0.5 m, and b) spacing of 15 m

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 spacing also increase, thus suggesting the adoption of appropriate and precise surveying tools together with a short spacing (high frequency of sampling) between spatial data points.

c7 Conclusions

"There are only two possible conclusions: if the result confirms the hypothesis, then you've just done a measure; if the result is contrary to the hypothesis, then you've made a discovery"
E.Fermi

In this final chapter, some conclusions have been organized separating remarks on each different parameters and techniques that have been analyzed through the document. About grouping process, two different identifying parameters have been studied: (1) the local deviation angle, and (2) the local curvature. Concerning the elements characterization, the analysis has been performed with three different alghoritms: (1) the Least Squares method, (2) the Landau technique, and (3) the Huber robust estimator. In the final part of these conclusions, some remarks on fitting results based on mobile mapping survey and GIS data are included.

7.1 About the local deviation angle

Some important conclusions can be drafted on the deviation angle as a possible parameter to identify the "transition points" between the geometrical components of the inspected road alignment.

In the literature (Li, Chitturi, Bill, & Noyce, 2012), the local deviation angle has been recognized and adopted as a parameter able to split up spatial data points into homogenous groups belonging to one of the three alignment elements (i.e., tangents, circular and spiral arcs). With this parameter it should be possible to pass from a unique differences-free point set to a series of groups on which successively operate some of the fitting techniques that have been explained in the previous paragraphs, obtaining desired parameters like curves radius, tangents azimuth and spiral scale factor.

This operation, called "grouping" or "element identification", requires that the adopted parameter must be sufficiently "sensitive" and "accurate" to distinguish the belonging to one of the alignment elements.

After this research study and looking at results, it is possible to state that **the local deviation angle is not a good parameter to perform the recognition and identification of design elements.**

Even if this is, let's say, a naïve or predictible conclusion, it represents a step forward in knowledge for this research field, since it addresses further research activities in different and more reliable algorithms.

In fact, it has been demonstrated that the deviation angle is strongly influenced by two survey parameters like the sampling frequency and the design geometry.

Firstly, it has been clearly shown that the space-based frequency is able to influence the local deviation angle and to disturb its detection up to levels where the signal itself is totally compromised and covered by a white random noise. Useful signals are obtained only with low frequencies (and then with a high spacing between the points): sampling frequencies of 0.5 m, 1 m, 2 m and 5 m results to be highly inappropriate to obtain a stable signal with low noise; instead, such results are sufficiently achieved with higher values of spacing between points (10 m, 15 m and 20 m).
This achievement could be useful and interesting in the survey planning stage, since it's a valid reason to set low

frequencies (in time or in space reference) and also it gives the opportunity to avoid choosing high sampling rate instruments (that generally are more expensive than other).

The deviation angle varies sensibly upon the spacing between points, not only in terms of "stability" of the signal, but also in quantitative terms, giving back different values under different sampling frequencies. This means also that the best point set with the best accuracy obtained with a mobile mapping survey will never give back a regular deviation angle profile, even along the same geometrical element. Due to unpredictable problems during the survey, like a cycle slip (a temporary loss of satellite signal due to obstacles) or an accidentally missing point, and also as a consequence of the natural instability of the car speed, the spacing between surveyed points can vary. This variation will give as a result a different local deviation angle value that will be variable even along the same curve. Theoretically, since the frequency in time reference is known, this phenomenon shouldn't be a big problem, but generally the survey is performed with a **time-based** fix frequency: this means that to obtain a stable frequency in space reference it should be necessary to keep a constant speed all along the survey. A "launch" and "stop" space before and after the surveyed area would be then necessary, but also small variations in speed can lead to big variations in space frequency and than in deviation angle signal.
Moreover, errors will appear due to slipping phenomena between vehicle tires and the road pavement.

In conclusion, it is possible to state that the deviation angle should not be considered for grouping identification.

7.2 LOCAL CURVATURE

The local curvature is also taken into consideration to perform grouping operation described in the previous paragraph (Wang, Zhencheng, & Uchimura, 2008).
Between the two methodologies adopted for the extraction of the local curvature, the fitting polynomial is able to give back best and more useful results.

In the thesis, it has been shown that the curvature obtained through the circumference on 3 points is too much sensitive to any inaccuracies in point positioning. Experimental results obtained do not allow to positively evaluate this technique: filtering with a moving average or particular treatments on sampling frequency are the only ways to partially reduce the error, which remains on unacceptable levels.

Furthermore, thanks to results obtained in the thesis, it can be stated that:

- there is a strong correlation between the average error obtained from the curvature extraction and the sampling frequency. Its effect reaches high levels, and the accuracy of the extracted curvature can vary up to three orders of magnitude only varying the space based sampling frequency from a minimum of 0.5 m to a maximum of 20 m;
- positioning accuracy varying from 1 cm to 40 cm does not seem to be heavily influent on the extracted curvature accuracy. In all cases its effect on results quality cannot be as effective as the sampling frequency;
- there is a correlation between the curvature accuracy and the maximum random error affecting the surveyed data points; with a third degree polynomial curve very high values of R^2 are obtained.

In conclusion, the local curvature obtained with the polynomial fitting is a good detecting parameter to perform the elements identification of road alignments. This variable is not heavily influenced by any survey parameter or setting (any variation could lead to a variable error noise that is common to every measurement process), and it is sufficiently tolerant with respect to local imprecisions or errors. The error transferred by the point positioning accuracy does not influence significantly the curvature and the adjustable fitting windows width is widely able to compensate any positioning error.

7.3 TANGENTS FITTING

Tangents characterization is the easiest recognition process in terms of results stability.

Independently from survey accuracy or from the space between them, **the quality of the fitting process is always good, giving back azimuth values with an error smaller than 1% of the target.** The fitting algorithms adopted in the current research faces with no difficulties every inaccuracy in point positioning and possible local defects, referring to quality boundaries investigated in this research.

In this work it has been confirmed that the extraction of tangent azimuth is sufficient to coherently perform a geometrical reconstruction of the road alignment: if all curves are well known (defined with radius and center coordinates) it's possible to extract all curve offsets (ΔR) and clothoids scale factors, obtaining a correct final alignment

The track continuity is also useful and sufficient to extract the length of all the geometrical element of the track, such as tangents, curves and clothoids, since their length will be defined by the curve termini (transitions points between the elements), and these points are univoquely determined by continuity conditions. This assumption is not valid only for the first and last element of the investigated alignment.

7.4 CURVES FITTING & BAD CONDITIONING

An interesting result obtained in this research is related to circular curves recognition. The circular arc is the most challenging element to be characterized and the most sensitive to factors like the point positioning accuracy, the sampling frequency and, most of all, the geometry of the problem, intended as the magnitude of radius and the length of the curve itself. When big radii are coupled with small curve lengths, a "bad conditioning situation" is present and problems in fitting process can occur.

About point positioning accuracy, it is possible to recognize a little reduction in fitting precision as long as the maximum random error increases too. Sampling frequency does not seem to have any sensible effect on fitting results: small variations are detectable between all frequencies, but these oscillations are generally random and it is not possible to observe a common trend.

It has been found that, for good fitting results, the width of the central angle is important. In fact, results have shown an high variability and inaccuracy that cannot be explained only with the the number of points, their positioning accuracy or sampling frequency. It has been observed an high connection between the central angle of the curve and the fitting results accuracy: the longer is the curve (or bigger is the central angle) the better will be the radius estimation independently form the algorithms used.

This relation has also been confirmed by a specific simulation in which a radii research has been performed on several curves with a known radius and of various lengths. Actually, it has been found that the radius estimated according to any fitting algorithm is smaller than the expected one: the difference between the real and the estimated value tends to reduce with the increase of the central angle covered by the curve length. It has been defined a threshold central angle which represents the minimum value needed to perform a good fitting. This angle also depends on the curve radius and with the point positioning accuracy.

Therefore, **the extraction of circular curve parameters (radius and center coordinates) are mainly affected by the geometrical conditions, and particularly from curve length and central angle.**

7.5 SURVEYED DATA FITTING RESULTS

Mobile Mapping and GIS based data have been also processed to evaluate local deviation angle, local curvature and the effectiveness of various fitting algorithms. It has been observed that the deviation angles and curvature are variable upon the quality of the data and their accuracy.

Good results in almost every evaluation have been obtained on GIS based data. Local curvatures as well as fitting results are good. Curvature is definitely a stable variable that can effectively leads to good grouping results since it generally follows the ideal curvature profile, with any polynomial degree and fitting window width. Fitting results on tangents are good (generally close to the real value with a deviation smaller than 1%). Curves are also correctly evaluated even if they suffer of bad conditioning situations.

Spatial data collected from mobile mapping techniques leads to good results under certain conditions, and generally suffers of error propagation effect. Data derived from the average of surveyed trajectories (IMU or GPS based) are more accurated than other alignments obtained from the image analysis. In fact, trajectory are directly obtained from positioning system onboard, while other set of data are extracted from the images captured by webcams. This means that the error inside the image extraction process is superposed to the error of the positioning system. This difference is relevant both for deviation angle and for curvature evaluation. Image analysis based tracks generally lead to deviation angles and curvature profiles that are practically unreadable, while fitting results are generally far away from the real values. An exception is represented by tangents that are generally well fitted with deviations lower than 10% from expected values).

7.6 FUTURE DEVELOPMENTS

From an operative point of view, it should be considered the opportunity to add to Mobile Mapping instrumentation a digital odometer, connected to a computer or a GPS to record the Computer time or the GPS one. This little add-on could really help to solve shifting problem on stations and give the possibility to compare any kind of data obtained during the survey or any successive elaboration. Thanks to this new instrument, there will be no more necessity to evaluate all the stations as the sum of successive chords, avoiding any further error.

Dealing with analysis algorithms development, the priority is the study and the application of a recognition algorithm that is able to find peaks inside an irregular signal. As it has been shown in the results, the curvature obtained with the polynomial fitting technique is a good parameter to identify the geometrical elements of an alignment, giving back a first separation between tangents, circular and transition curves. To identify all the transition point from a single element to another one, a new algorithm is needed. This algorithm should analyze the extracted curvature signal, recognize maximum and minimum peaks on extended intervals (thus avoiding local peaks), and then define some transition stations.

Bad geometrical conditions situation needs to be investigated deeper since road curves are often inside these conditions. In fact, where big radii and small lengths are present, fitting problems araise.

It will be also really promising to continue the testing of fitting algorithms towards new robust model estimators, like Mallows or Shweppe (Huber & Ronchetti, 2009) (ref.), that seems to be really effective on leverage points (2-dimension outliers). These outliers can be common due to unpredictable inaccuracies or errors during the survey stage, like cycle slips or bad satellite coverage (that is whe visible satellites are too few).