Reports on Geodesy and Geoinformatics, 2025, Vol. 120, pp. 67–74

DOI: 10.2478/rgg-2025-0017
Received: 16 July 2025 / Accepted: 13 October 2025
Published online: 12 November 2025

ORIGINAL ARTICLE

Laser scanning data processed using M_{split} estimation and sliding window algorithm

Robert Duchnowski [1] and Patrycja Wyszkowska [1*]

[1]Department of Geodesy, Institute of Geodesy and Civil Engineering, Faculty of Geoengineering, University of Warmia and Mazury in Olsztyn, Oczapowskiego 2, 10-719 Olsztyn, Poland

*patrycja.wyszkowska@uwm.edu.pl

Abstract

Laser scanning systems are modern measurement techniques generating large datasets. Observations, usually collected as a point cloud, present the general results that can be visualized using specialized software. While the final effect might be impressive from a visualization point of view, it is inconvenient for modeling or extracting detailed information about, for example, terrain, buildings, engineering structures, and deformations. Therefore, data from laser scanning systems require post-processing using several methods reflecting different purposes or data processing stages: data segmentation, modeling, and filtration. M_{split} estimation is one of the methods that has proved its effectiveness in laser scanning data processing and determination of terrain profiles, deformation, or building shapes. Processing the complete datasets tends to only yield often inadequate results when high-class computers are used, and it is time-consuming. Therefore, datasets tend to remain segmented. This paper explores a range of several types of segmentation methods that can be used in M_{split} estimation. It presents profile determination when data cut out from the original point cloud are divided into intervals of the same length, or the sliding window algorithm is applied. In comparison, the given examples show that the latter approach can provide more reliable results. The application of the sliding window algorithm entails having to make assumptions concerning estimation parameters. The paper offers valuable guidance about both the width of the window and the slide size.

Key words: sliding window algorithm, M_{split} estimation, laser scanning

1 Introduction

The advanced measurement techniques contain novel systems of data acquisition that often generate large datasets. Light Detection and Ranging (LiDAR) is one of such systems that has gained popularity and is applied to solve many engineering problems, including surveying (Yang et al., 2017; Janicka et al., 2020; Wyszkowska et al., 2020), geomatics (Lian and Hu, 2017; Zhao et al., 2019; Błaszczak-Bąk et al., 2020), civil engineering (Wang and Hsu, 2007; Cabaleiro et al., 2015; Błaszczak-Bąk et al., 2020; Wyszkowska and Duchnowski, 2022), geosciences (Spaete et al., 2010), archeology (Rodríguez-Gonzálvez et al., 2017), forestry (Crespo-Peremarch et al., 2018; Arslan et al., 2021). LiDAR technology encompasses three main types: Terrestrial Laser Scanning (TLS), Airborne Laser Scanning (ALS), and Mobile Laser Scanning (MLS). The measurements are usually grouped and presented in a set named a point cloud (usually in 3D space), regardless of the technique applied. A point cloud can provide valuable information about the scanned object, for example, a building, or any engineering structures, terrain, treetops, vegetation cover, or excavation. However, the raw data usually only provide general information. To obtain more detailed descriptions, data must undergo a process consisting of several stages, including registration, segmentation, data cleaning, filtration, modelling, or estimation of parameters of geometrical primitives such as surfaces, profiles (e.g., Forlani and Nardinocchi, 2007; Tóvári and Pfeifer, 2005; Nguyen and Le, 2013; Błaszczak-Bąk et al., 2015). Processing complete sets is sufficient and easy to perform only when the relevant point clouds include a relatively small number of measurements and the object under study is simple. In other cases, processing the whole sets is significantly time-consuming and requires a high-class computer. Moreover, when processing

Publisher: De Gruyter

the full sets, more detailed information about the object may be lost. Therefore, the point clouds are usually divided into subsets to make the processing more efficient. Observation segmentation concerns usually complex objects like buildings, bridges, etc., where each construction element might be modelled separately. However, segmentation might also be applied to perform a more detailed analysis to obtain a more comprehensive model.

Data processing also faces the problem of the choice of estimation or modelling method. There is no doubt that the least-squares method (LS estimation) is the most common. However, other approaches are also used, including M-estimation, robust estimation, random sample consensus (RANSAC), and M_{split} estimation (e.g., Carrilho et al., 2018; Zhao et al., 2019). The latter method is relatively new and has found interesting applications to surveying engineering problems: finding unstable points in global navigation satellite systems (GNSS) networks (Banimostafavi et al., 2023), deformation analysis (e.g., Zienkiewicz, 2015; Duchnowski and Wyszkowska, 2022b; Pleterski et al., 2025), similarity transformation (Nowel, 2018; Zhang et al., 2023), estimation with errors-in-variables (EIV) models (Wiśniewski, 2022), data processing with the point cloud spatial expansion (PCSE) algorithm (Zienkiewicz and Dąbrowski, 2023), heterogeneous data fusion (Tao et al., 2024, 2025), direct identification of outliers (Li et al., 2013), robust estimation (Wyszkowska and Duchnowski, 2022, 2024b), and coordinate transformation (Janicka and Rapinski, 2013). M_{split} estimation, in different variants, was also successfully applied to process LIDAR data in the following problems: displacement analysis (Janicka et al., 2020, 2023; Wyszkowska and Duchnowski, 2022), detection and analysis of engineering structures (Janicka and Rapinski, 2013; Janowski, 2018; Dąbrowski and Zienkiewicz, 2022; Wyszkowska and Duchnowski, 2024b), terrain modeling (Błaszczak-Bąk et al., 2015; Wyszkowska et al., 2020). The wide range of M_{split} estimation applications mostly stems from its unique feature, namely, it can estimate parameters within the split functional model. In practice, it means that by processing one observation set, one can determine two (or more) versions of parameters (in some sense, the method is similar to cluster analysis; however, the algorithms, assumptions, and data processing are dissimilar – the observation sets are not divided into clusters processed separately, in fact) (Wiśniewski, 2009, 2010). This unique feature also enables applying the method as an alternative to robust estimation, which might overperform the classical approaches, including M-estimation or R-estimation (Duchnowski and Wyszkowska, 2022a; Wyszkowska and Duchnowski, 2022, 2024a).

Estimating two or more variants of functional model parameters is the main advantage of M_{split} estimation. It is a unique property as other methods have to process subsets separated in a specific way to provide comparable results. Another advantage of the method is its robustness against outliers. From the theoretical point of view, the basic M_{split} estimation variants cannot be classified as robust against outliers (Duchnowski and Wiśniewski, 2019; Duchnowski and Wyszkowska, 2023); however, they can be used as alternatives to robust methods like M-estimation. In such an application, the method should generate regular observations from outliers. During the estimation process, location parameters of both groups are estimated (Duchnowski and Wyszkowska, 2022a). Notably, robust variants of M_{split} estimation have also been derived (Wyszkowska and Duchnowski, 2022, 2024a). Crucially, in robust applications, M_{split} estimation can withstand a significantly high percentage of outliers (even more than 50%), which conventional robust methods fail to deliver (Wyszkowska and Duchnowski, 2022; Duchnowski and Wyszkowska, 2023). Like every estimation method, M_{split} estimation has some disadvantages: Its algorithms are more complex than the algorithms of M-estimation; in some applications, it is also important to select sufficiently accurate starting points, which requires some experience from analysts (Wyszkowska and Duchnowski, 2019, 2020). Nevertheless, the advantages of this method outweigh the limitations in many surveying applications.

This paper addresses processing LiDAR data by applying M_{split} estimation. It explores acquiring information from a point cloud and examines several scenarios in which the observations are processed. The data subsets are cut out from the whole point cloud and become the basis for modelling characteristic elements of the object. Each observation set can be processed as a whole or divided into several subsets. Another option is the application of the sliding window algorithm (e.g., Wang et al., 2016; Li et al., 2018). Wyszkowska and Duchnowski (2025) applied the method to M_{split} estimation; however, such an approach has never been examined in detail. Therefore, this paper lists advantages and disadvantages of the sliding window algorithm in relation to processing the whole observation set or subsets mentioned when M_{split} estimation is used.

The paper is organized in the following way. The Section 2 presents the foundations of M_{split} estimation in two basic variants. Section 3 summarizes examples of modelling the wall profile and the wall edge for different observation sets (also disturbed by outliers). The final section discusses the results and presents conclusions regarding applying the sliding window algorithm to M_{split} estimation method in LiDAR data processing.

2 M_{split} estimation

The main assumption of M_{split} estimation is that the observation set is an unknown mixture of realizations of at least two different random variables (the observation set consists, in fact, of at least two subsets differing in location parameters; however, the set division stays unknown). Therefore, the original functional model of observations is split into at least two competitive models as follows:

$$\mathbf{y} = \mathbf{AX} + \mathbf{v} \Rightarrow \begin{cases} \mathbf{y} = \mathbf{AX}_{(1)} + \mathbf{v}_{(1)} \\ \mathbf{y} = \mathbf{AX}_{(2)} + \mathbf{v}_{(2)} \end{cases} \tag{1}$$

where: $\mathbf{y}$ – observation vector, $\mathbf{A}$ – full column rank coefficient matrix, $\mathbf{X}$ – parameter vector, $\mathbf{v}$ – observation error vector, $\mathbf{X}_{(1)}$, and $\mathbf{X}_{(2)}$ – competitive versions of parameter vector $\mathbf{X}$, $\mathbf{v}_{(1)}$ and $\mathbf{v}_{(2)}$ – competitive versions of observation error vector $\mathbf{v}$. The first and basic variant of M_{split} estimation is the squared M_{split} estimation (SMS), which was proposed in Wiśniewski (2009). The objective function $\varphi\left(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}\right)$, the influence functions $\psi_{(l)}\left(v_{i(1)}, v_{i(2)}\right)$, and the weight functions $w_{(l)}\left(v_{i(1)}, v_{i(2)}\right)$ (where $l = 1$ or 2) are defined in the following way:

$$\begin{cases} \varphi\left(\mathbf{X}_{(1)}, \mathbf{X}_{(2)}\right) = \sum_{i=1}^{n} \rho\left(v_{i(1)}, v_{i(2)}\right) = \sum_{i=1}^{n} v_{i(1)}^2 v_{i(2)}^2 \\ \psi_{(1)}\left(v_{i(1)}, v_{i(2)}\right) = 2v_{i(1)} v_{i(2)}^2 \\ \psi_{(2)}\left(v_{i(1)}, v_{i(2)}\right) = 2v_{i(1)}^2 v_{i(2)} \\ w_{(1)}\left(v_{i(1)}, v_{i(2)}\right) = v_{i(2)}^2 \\ w_{(2)}\left(v_{i(1)}, v_{i(2)}\right) = v_{i(1)}^2 \end{cases} \tag{2}$$

The second variant, called the absolute M_{split} estimation (AMS), was proposed in Wyszkowska and Duchnowski (2019), and its main functions are as follows:

$$\begin{cases} \varphi\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right) = \sum_{i=1}^{n}\rho\left(v_{i(1)},v_{i(2)}\right) = \sum_{i=1}^{n}\left|v_{i(1)}\right|\left|v_{i(2)}\right| \\ \psi_{(1)}\left(v_{i(1)},v_{i(2)}\right) = \begin{cases} -\left|v_{i(2)}\right| & \text{for } v_{i(1)} < 0 \\ \left|v_{i(2)}\right| & \text{for } v_{i(1)} > 0 \end{cases} \\ \psi_{(2)}\left(v_{i(1)},v_{i(2)}\right) = \begin{cases} -\left|v_{i(1)}\right| & \text{for } v_{i(2)} < 0 \\ \left|v_{i(1)}\right| & \text{for } v_{i(2)} > 0 \end{cases} \\ w_{(1)}\left(v_{i(1)},v_{i(2)}\right) = \begin{cases} -\frac{\left|v_{i(2)}\right|}{2v_{i(1)}} & \text{for } v_{i(1)} < 0 \\ \frac{\left|v_{i(2)}\right|}{2v_{i(1)}} & \text{for } v_{i(1)} > 0 \end{cases} \\ w_{(2)}\left(v_{i(1)},v_{i(2)}\right) = \begin{cases} -\frac{\left|v_{i(1)}\right|}{2v_{i(2)}} & \text{for } v_{i(2)} < 0 \\ \frac{\left|v_{i(1)}\right|}{2v_{i(2)}} & \text{for } v_{i(2)} > 0 \end{cases} \end{cases} \tag{3}$$

The parameters of the functional models (1) are estimated in the iterative process. SMS estimation uses the traditional iterative process proposed in Wiśniewski (2009):

$$\begin{aligned} \mathbf{X}_{(1)}^{j} &= \mathbf{X}_{(1)}^{j-1} + d\mathbf{X}_{(1)}^{j} = \mathbf{X}_{(1)}^{j-1} - \left[\mathbf{H}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right) \\ \mathbf{X}_{(2)}^{j} &= \mathbf{X}_{(2)}^{j-1} + d\mathbf{X}_{(2)}^{j} = \mathbf{X}_{(2)}^{j-1} - \left[\mathbf{H}_{(2)}\left(\mathbf{X}_{(1)}^{j},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{j},\mathbf{X}_{(2)}^{j-1}\right) \end{aligned} \tag{4}$$

where: $d\mathbf{X}_{(l)}$ – increment to parameter vector, $\mathbf{H}_{(l)}\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)$ – Hessians, $\mathbf{g}_{(l)}\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)$ – gradients. The Hessians and gradients are defined as follows:

$$\begin{aligned} \mathbf{H}_{(1)}\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right) &= \frac{\partial^2\varphi\left(\mathbf{y};\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)}{\partial\mathbf{X}_{(1)}\partial\mathbf{X}_{(1)}^{T}} = 2\mathbf{A}^{T}\mathbf{w}_{(1)}\left(\mathbf{v}_{(1)},\mathbf{v}_{(2)}\right)\mathbf{A} \\ \mathbf{H}_{(2)}\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right) &= \frac{\partial^2\varphi\left(\mathbf{y};\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)}{\partial\mathbf{X}_{(2)}\partial\mathbf{X}_{(2)}^{T}} = 2\mathbf{A}^{T}\mathbf{w}_{(2)}\left(\mathbf{v}_{(1)},\mathbf{v}_{(2)}\right)\mathbf{A} \end{aligned} \tag{5}$$

$$\begin{aligned} \mathbf{g}_{(1)}\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right) &= \left[\frac{\partial\varphi\left(\mathbf{y};\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)}{\partial\mathbf{X}_{(1)}}\right]^{T} = -2\mathbf{A}^{T}\mathbf{w}_{(1)}\left(\mathbf{v}_{(1)},\mathbf{v}_{(2)}\right)\mathbf{v}_{(1)} \\ \mathbf{g}_{(2)}\left(\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right) &= \left[\frac{\partial\varphi\left(\mathbf{y};\mathbf{X}_{(1)},\mathbf{X}_{(2)}\right)}{\partial\mathbf{X}_{(2)}}\right]^{T} = -2\mathbf{A}^{T}\mathbf{w}_{(2)}\left(\mathbf{v}_{(1)},\mathbf{v}_{(2)}\right)\mathbf{v}_{(2)} \end{aligned} \tag{6}$$

where the matrices of the weight functions are recomputed in each iterative step, in the following way:

$$\begin{aligned} \mathbf{w}_{(1)}\left(\mathbf{v}_{(1)},\mathbf{v}_{(2)}\right) &= \operatorname{diag}\left[w_{(1)}\left(v_{1(1)},v_{1(2)}\right),\ldots,w_{(1)}\left(v_{n(1)},v_{n(2)}\right)\right] \\ \mathbf{w}_{(2)}\left(\mathbf{v}_{(1)},\mathbf{v}_{(2)}\right) &= \operatorname{diag}\left[w_{(2)}\left(v_{1(1)},v_{1(2)}\right),\ldots,w_{(2)}\left(v_{n(1)},v_{n(2)}\right)\right] \end{aligned} \tag{7}$$

where: diag $(\circ)$ – diagonal matrix. Therefore, the weight matrices depend on the weight functions and the values of the errors $v_{i(1)}$ either $v_{i(2)}$. This method is called mutual cross-weighting. The iterative process ends for such j for which the necessary conditions $\mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right) = 0$ and $\mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right) = 0$ are satisfied. Hence $\hat{\mathbf{X}}_{(1)} = \mathbf{X}_{(1)}^{j} = \mathbf{X}_{(1)}^{j-1}$ and $\hat{\mathbf{X}}_{(2)} = \mathbf{X}_{(2)}^{j} = \mathbf{X}_{(2)}^{j-1}$ (Wiśniewski, 2009; Wyszkowska and Duchnowski, 2019).

The traditional process is not applicable where the mutual cross-weighting is not applied, for example, in AMS estimation (see Equation 3), in which case a parallel iterative process proposed

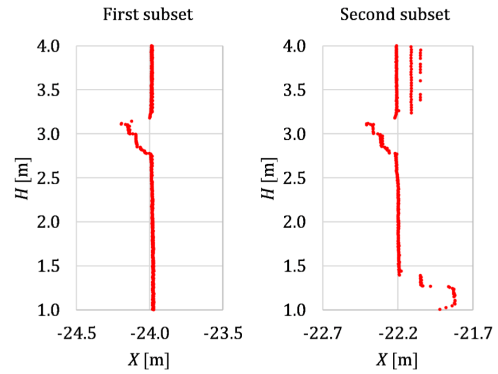

Figure 1. Data sets of the wall with cornice

in Wyszkowska and Duchnowski (2019) is applicable:

$$\begin{aligned} \mathbf{X}_{(1)}^{j} &= \mathbf{X}_{(1)}^{j-1} + d\mathbf{X}_{(1)}^{j} = \mathbf{X}_{(1)}^{j-1} - \left[\mathbf{H}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(1)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right) \\ \mathbf{X}_{(2)}^{j} &= \mathbf{X}_{(2)}^{j-1} + d\mathbf{X}_{(2)}^{j} = \mathbf{X}_{(2)}^{j-1} - \left[\mathbf{H}_{(2)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right)\right]^{-1}\mathbf{g}_{(2)}\left(\mathbf{X}_{(1)}^{j-1},\mathbf{X}_{(2)}^{j-1}\right) \end{aligned} \tag{8}$$

The gradients and Hessians are defined as in Equations (5) and (6), and the criterion for ending the iterative process is the same as in the traditional iterative process.

3 Processing example TLS data

Examining the presented scenarios is performed on the example object, a building placed at the campus of the University of Warmia and Mazury in Olsztyn, Poland. The object was scanned using Leica ScanStation C10 (terrestrial laser scanner). Several data subsets were extracted from the whole point cloud to determine wall profiles or edges. The subsets were designed to demonstrate the basic applications of M_{split} estimation, namely: the natural one, estimating competitive versions of the parameters from the functional model (1), or application of the method as an alternative to the conventional robust estimations.

3.1 Extracting wall profiles

The section explores two observation sets (presented in Figure 1) created to determine the wall profiles. The first one contains measurements of the wall profile with a cornice around the middle. The second subset consists of measurements of the profile parallel to the first one, and it contains some outliers resulting from measuring the window recess and the room behind (in the upper part). The observations in the lower part of the subsets describe the door recess; they should be regarded as regular ones, but they disturb the profile's linearity.

Here, five scenarios of data processing are proposed:

- Scenario A – each profile of a length of 3 m processed as a whole,
- Scenario B – 0.2 m intervals (from 1.1 m to 1.3 m, from 1.3 m to 1.5 m, etc.),
- Scenario C – 0.1 m intervals (from 1.05 m to 1.15 m, from 1.15 m to 1.25 m, etc.),
- Scenario D – 0.4 m sliding windows and the window slide of 0.2 m (from 1.0 m to 1.4 m, from 1.2 m to 1.6 m, etc.),
- Scenario E – 0.2 m sliding windows and the window slide of 0.1 m (from 1.0 m to 1.2 m, from 1.1 m to 1.3 m, etc.).

In all variants, the parameters from the models of Equation (1) are regarded as the parameters of the linear function (the first-degree polynomial). The estimated profiles from different scenarios are presented in Figures 2 and 3, which also present profiles result-

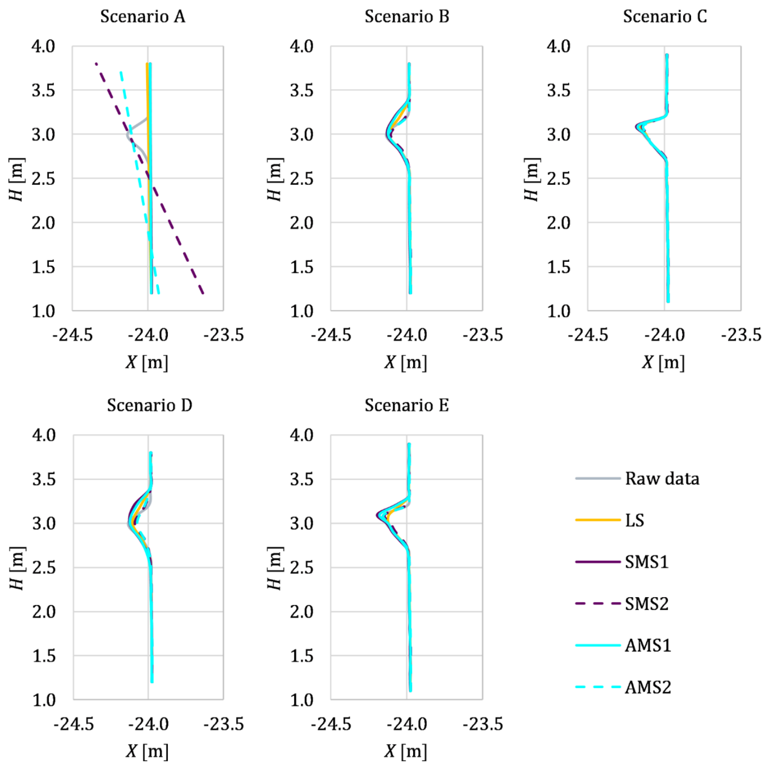

Figure 2. Profile estimated from the first subset across different scenarios (solid lines – final solutions, dashed lines – alternative solutions)

ing from raw data interpolation. In the case of M_{split} estimation, two possible solutions are presented; the solid lines represent the first solutions (SMS1 or AMS1), whereas the dashed lines represent the second solutions (SMS2 or AMS2). SMS1 and AMS1 solutions are the ones that describe the wall profiles (the others might be regarded as describing outlying observations).

A simple graphical analysis yields conclusions: First, processing the whole set might provide satisfactory results when no outliers occur. However, even then some detailed information is lost (see the first subset and the cornice). Second, as expected, the shorter the intervals (or the sliding window and the slide), the more detailed the profile can be obtained. However, it is hard to decide which method of dividing the whole observation set provides better results. On the other hand, one can notice that M_{split} estimations outperform LS estimation, especially in the case of the second set. Comparing the results obtained for that case in Scenarios C and D with those of Scenarios E and F, one can see that the sliding window algorithm provides slightly better results in the upper profile part. For more detailed and informative analyses, a reference is necessary, for example, a line or a surface determined with much higher accuracy. This is not the case in that example because of the shape of the profiles and the measurements performed. Therefore, we introduce another example where such a surface can be defined.

3.2 Determining the wall edges

Two additional observation sets were created to determine the wall edges. Since those parts of the object were scanned with the highest resolution, it is possible to create the examined sets and the reference ones- four times bigger and with no outliers. Therefore, the reference sets can determine the wall edges more precisely. All sets are presented in Figures 4 and 5.

The data processing in the first case is a natural application of M_{split} estimation. Generally, that set consists of measurements of both sides of the corner, and the estimation method mentioned allows us to estimate parameters of two planes modelling the walls from both sides in one iterative process. From such estimated parameters, one can compute the line that is the intersection of the planes, and it is the model of the wall edge, meanwhile, the reference wall edge is computed from the separate reference subsets using LS estimation. Each subset in question consists of measurements of only one side of the corner; hence, it allows for the modelling of that corner side.

The following are scenarios of data processing:

- Scenario A – data processed as a whole,
- Scenario B – 0.5 m intervals in relation to Z-axis (for Z from 3.595 m to 4.095 m, from 4.095 m to 4.595 m, etc., with the last interval larger),
- Scenario C – 0.25 m intervals in relation to Z-axis (for Z from 3.595 m to 3.845 m, from 3.845 m to 4.095 m, etc., with the last interval larger),
- Scenario D – 0.5 m sliding windows in relation to Z-axis and the window slide of 0.25 m (for Z from 3.595 m to 4.095 m, from 3.845 m to 3.435 m, etc., with the last sliding window is larger),
- Scenario E – 0.25 m sliding windows in relation to Z-axis and the window slide of 0.125 m (for Z from 3.595 m to 3.845 m, from 3.720 m to 3.970 m, etc., with the last sliding window larger).

To compare the precision of the obtained wall edge from different scenarios, the root-mean-square deviation (RMSD) can be applied, following a determination of the reference edge in each interval using the reference set, then a determination of the point of intersection of that wall edge with the horizontal plane at the chosen heights. The same should be done for the examined sets. The distance between such points, d_i, is the base to compute RMSD using the following formula:

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} \tag{9}$$

where n is the number of distances between the points determined. Crucially, computed RMSDs do not present the real accuracy of

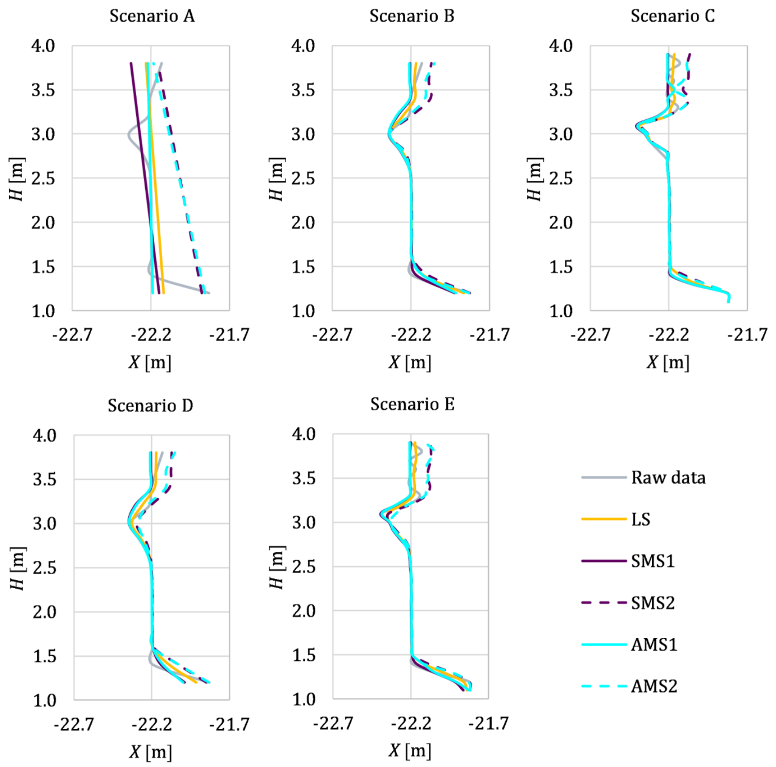

Figure 3. Profile estimated from the second subset across different scenarios (solid lines – final solutions, dashed lines – alternative solutions)

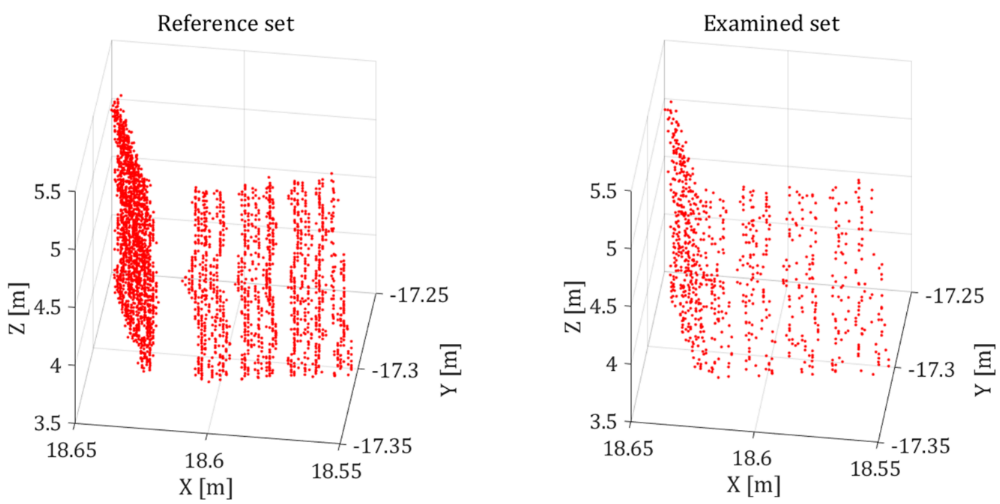

Figure 4. TLS point clouds for wall edge determination (the first case)

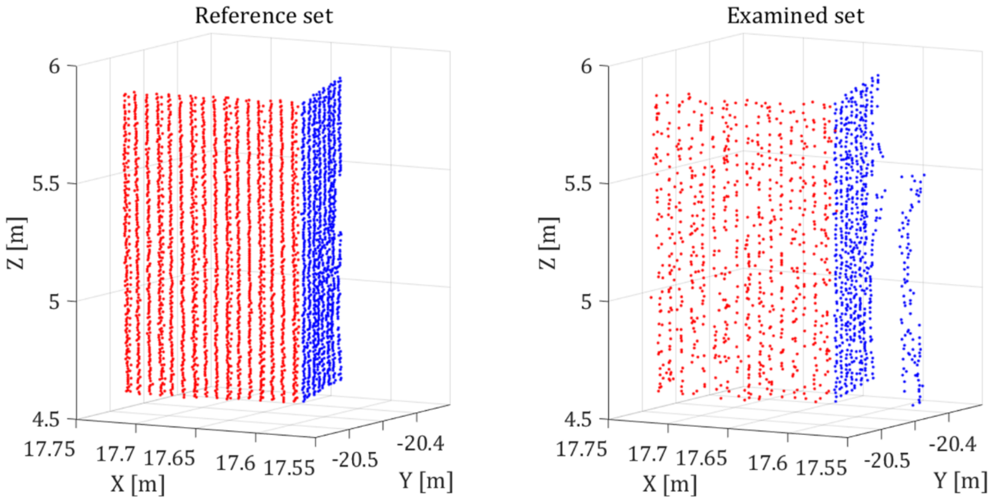

Figure 5. TLS point clouds for wall edge determination (the second case)

Table 1. RMSDs [mm] for different scenarios (the first case)

Scenario	SMS	AMS
A	2.57	2.13
B	1.25	0.87
C	1.43	1.17
D	1.41	0.91
E	1.24	0.89

Table 2. RMSDs [mm] for different scenarios (the second case)

Scenario	SMS	AMS
A	0.92	0.74
B	0.57	0.33
C	0.72	0.66
D	0.56	0.32
E	0.63	0.48

the data processing methods; however, they can be used to make comparative analyses. Table 1 presents RMSDs obtained for all scenarios in the first case. One can say that processing data divided into subsets provides better results than processing the complete set. At this point it remains unclear which strategy is more reliable: the sliding window algorithm or dividing the complete set into intervals in relation to Z-axis. With narrower intervals, the results are of a similar accuracy (compare scenarios B and C). On the other hand, assuming the narrower sliding window, one more accurate results can be obtained.

In the second case, the examined set contains outliers (measurements of the rain gutter; see right panel of Figure 5). Therefore, the measurements concerning the wall's left or right sides are processed separately. It means that M_{split} estimation is here a viable alternative for the robust procedure.

The following are scenarios of data processing:

- Scenario A – data processed as a whole,
- Scenario B – 0.4 m intervals in relation to Z-axis (for Z from 4.566 m to 4.966 m, from 4.966 m to 5.366 m, etc., with the last interval larger),
- Scenario C – 0.25 m intervals in relation to Z-axis (for Z from 4.566 m to 4.816 m, from 4.816 m to 5.066 m, etc., with the last interval larger),
- Scenario D – 0.4 m sliding windows in relation to Z-axis and the window slide of 0.2 m (for Z from 4.566 m to 4.966 m, from 4.766 m to 5.166 m, etc., with the last sliding window larger),
- Scenario E – 0.25 m sliding windows in relation to Z-axis and the window slide of 0.125 m (for Z from 4.566 m to 4.816 m, from 4.691 m to 4.941 m, etc., with the last sliding window larger).

To compare results for those scenarios, RMSDs computed as in the previous example can be used. The results are presented in Table 2.

Processing the complete set at that time provided the least accurate results. The sliding window algorithm and the processing of the subsets provide similar results; however, the former method proved more accurate. Interestingly, narrowing the intervals (or the sliding windows) deteriorates the precision of the wall edge determination.

4 Discussion and conclusions

This paper investigates LiDAR data processing using M_{split} estimation for modeling engineering structures across various scenarios. A central distinction among the processing approaches lies in how the data are partitioned. Specifically, the dataset may be processed as a whole, divided into mutually exclusive subsets, or analyzed using a sliding window algorithm.

The examples presented demonstrate that processing the entire dataset without segmentation can yield satisfactory results only when the modelled surface or profile is geometrically simple – such as those that can be effectively described by polynomials or other basic functions. In such cases, the data typically contains few outliers, allowing for accurate and efficient modelling. However, when these conditions are not met – particularly in the presence of complex geometries or significant noise – this approach may lead to unsatisfactory or inaccurate outcomes. Therefore, selecting an appropriate data partitioning strategy is crucial for ensuring robust and reliable modelling performance in diverse engineering contexts.

Comparing the approaches based on processing the subsets or the sliding window method provides ambiguous outcomes. In the first example, the results of both approaches seem almost identical; for both methods, assuming narrower intervals or sliding windows, respectively, yields more precise results. In the second example, the conclusions are different. The more accurate models usually come from wider intervals or sliding windows, which reflects the characteristics of the observation set under investigation and the object model. At that time, one tries to model the wall edge; hence, 3D sets are processed. Therefore, narrowing intervals or sliding windows might result in processing subsets whose dimensions along three axes are significantly different (the subsets become elongated along the wall faces). In that case, modelling the planes becomes more sensitive to the local fluctuations of the wall surfaces. The second example shows that the sliding window algorithm mostly provides better results.

Considering the results presented, one can suggest the application of the sliding window algorithm when modelling long engineering structures, such as profiles or wall edges. However, it is more time-consuming than processing data in mutually exclusive subsets; this approach can provide more reliable and accurate results. The outstanding problem is the selection of the size of the sliding window and the slide itself (similarly to the size of subsets). As shown, this size should reflect the modelled structure to describe even small characteristic elements. On the other hand, a reduction in size may compromise the final results. Therefore, the sliding window size should be assumed for each observation set separately, depending on the set size and its resolution. Finally, the paper also shows that processing the complete observation sets only yields acceptable accuracy when the shape under study is relatively simple and easy to model.

References

Arslan, A. E., Erten, E., and Inan, M. (2021). A comparative study for obtaining effective Leaf Area Index from single Terrestrial Laser Scans by removal of wood material. *Measurement*, 178:109262, doi:10.1016/j.measurement.2021.109262.

Banimostafavi, Z., Sharifi, M. A., and Farzaneh, S. (2023). Evaluation of unstable points detection methods in geodetic GNSS-based networks. *Iranian Journal of Geophysics*, 16(4):175–192, doi:10.30499/ijg.2023.350587.1441.

Błaszczak-Bąk, W., Janowski, A., Kamiński, W., and Rapiński, J. (2015). Application of the Msplit method for filtering airborne laser scanning data-sets to estimate digital terrain models. *International Journal of Remote Sensing*, 36(9):2421–2437, doi:10.1080/01431161.2015.1041617.

Błaszczak-Bąk, W., Suchocki, C., Janicka, J., Dumalski, A., Duchnowski, R., and Sobieraj-Żłobińska, A. (2020). Automatic Threat Detection for Historic Buildings in Dark Places Based on the Modified OptD Method. *ISPRS International Journal of Geo-Information*, 9(2):123, doi:10.3390/ijgi9020123.

Cabaleiro, M., Riveiro, B., Arias, P., and Caamaño, J. (2015). Algorithm for beam deformation modeling from LiDAR data. *Measurement*, 76:20–31, doi:10.1016/j.measurement.2015.08.023.

Carrilho, A. C., Galo, M., and Santos, R. C. (2018). Statistical Outlier Detection Method For Airborne LiDAR Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII–1:87–92, doi:10.5194/isprs-archives-xlii-1-87-2018.

Crespo-Peremarch, P., Tompalski, P., Coops, N. C., and Ruiz, L. A. (2018). Characterizing understory vegetation in Mediterranean forests using full-waveform airborne laser scanning data. *Remote Sensing of Environment*, 217:400–413, doi:10.1016/j.rse.2018.08.033.

Duchnowski, R. and Wiśniewski, Z. (2019). Robustness of Msplit(q) estimation: A theoretical approach. *Studia Geophysica et Geodaetica*, 63(3):390–417, doi:10.1007/s11200-018-0548-x.

Duchnowski, R. and Wyszkowska, P. (2022a). Absolute Msplit estimation as an alternative for robust M-estimation. *Advances in Geodesy and Geoinformation*, page 17–17, doi:10.24425/gac.2022.141170.

Duchnowski, R. and Wyszkowska, P. (2022b). Unstable Object Points during Measurements—Deformation Analysis Based on Pseudo Epoch Approach. *Sensors*, 22(23):9030, doi:10.3390/s22239030.

Duchnowski, R. and Wyszkowska, P. (2023). Tolerance for Growing Errors of Observations as a Measure Describing Global Robustness of Msplit Estimation and Providing New Information on Other Methods. *Journal of Surveying Engineering*, 149(4), doi:10.1061/jsued2.sueng-1451.

Dąbrowski, P. S. and Zienkiewicz, M. H. (2022). Impact of cross-section centers estimation on the accuracy of the point cloud spatial expansion using robust M-estimation and Monte Carlo simulation. *Measurement*, 189:110436, doi:10.1016/j.measurement.2021.110436.

Forlani, G. and Nardinocchi, C. (2007). Adaptive filtering of aerial laser scanning data. In *Workshop on Laser Scanning 2007 and SilviLaser 2007, 12-14 September 2007, Espoo, Finland*, volume 36, pages 130–35.

Janicka, J. and Rapinski, J. (2013). Msplit transformation of coordinates. *Survey Review*, 45(331):269–274, doi:10.1179/003962613x13726661625708.

Janicka, J., Rapinski, J., and Błaszczak-Bąk, W. (2023). Orthogonal Msplit Estimation for Consequence Disaster Analysis. *Remote Sensing*, 15(2):421, doi:10.3390/rs15020421.

Janicka, J., Rapiński, J., Błaszczak-Bąk, W., and Suchocki, C. (2020). Application of the Msplit Estimation Method in the Detection and Dimensioning of the Displacement of Adjacent Planes. *Remote Sensing*, 12(19):3203, doi:10.3390/rs12193203.

Janowski, A. (2018). The circle object detection with the use of Msplit estimation. *E3S Web of Conferences*, 26:00014, doi:10.1051/e3sconf/20182600014.

Li, J., Wang, A., and Xinyuan, W. (2013). Msplit estimate the relationship between LS and its application in gross error detection. *Mine Surveying*, 2:57–59.

Li, J., Xiong, B., Biljecki, F., and Schrotter, G. (2018). A sliding window method for detecting corners of openings from terrestrial LiDAr data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42(4/W10):97–103, doi:10.5194/isprs-archives-XLII-4-W10-97-2018.

Lian, X. and Hu, H. (2017). Terrestrial laser scanning monitoring and spatial analysis of ground disaster in Gaoyang coal mine in Shanxi, China: a technical note. *Environmental Earth Sciences*, 76(7), doi:10.1007/s12665-017-6609-6.

Nguyen, A. and Le, B. (2013). 3D point cloud segmentation: A survey. In *2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, page 225–230. IEEE, doi:10.1109/ram.2013.6758588.

Nowel, K. (2018). Squared Msplit(q) S-transformation of control network deformations. *Journal of Geodesy*, 93(7):1025–1044, doi:10.1007/s00190-018-1221-4.

Pleterski, Ž., Ambrožič, T., Mulahusić, A., Tuno, N., Topoljak, J., Hajdar, A., Hamzić, A., Đidelija, M., Kulo, N., Rak, G., Marjetič, A., and Kregar, K. (2025). Squared Msplit estimation in deformation analysis – 2D geodetic network case study. *Geodetski vestnik*, 69(2):115–147, doi:10.15292/geodetski-vestnik.2025.02.115-147.

Rodríguez-Gonzálvez, P., Jiménez Fernández-Palacios, B., Muñoz-Nieto, A., Arias-Sanchez, P., and Gonzalez-Aguilera, D. (2017). Mobile LiDAR System: New Possibilities for the Documentation and Dissemination of Large Cultural Heritage Sites. *Remote Sensing*, 9(3):189, doi:10.3390/rs9030189.

Spaete, L. P., Glenn, N. F., Derryberry, D. R., Sankey, T. T., Mitchell, J. J., and Hardegree, S. P. (2010). Vegetation and slope effects on accuracy of a LiDAR-derived DEM in the sagebrush steppe. *Remote Sensing Letters*, 2(4):317–326, doi:10.1080/01431161.2010.515267.

Tao, Y., Li, X., Chen, H., and Yang, J. (2024). Solution for the Robust Estimation of Heterogeneous Data Fusion Based on Classification Estimation. *Journal of Surveying Engineering*, 150(3), doi:10.1061/jsued2.sueng-1492.

Tao, Y., Su, M., Xu, Z., and Chen, H. (2025). Solution for heterogeneous data fusion based on autonomous classification. *Measurement*, 243:116326, doi:10.1016/j.measurement.2024.116326.

Tóvári, D. and Pfeifer, N. (2005). Segmentation based robust interpolation-a new approach to laser data filtering. *Interna-*

tional Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/19):79–84.

Wang, A., Li, C., Liu, Y., Zhuang, Y., Bu, C., and Xiao, J. (2016). Laser-Based Online Sliding-Window Approach for UAV Loop-Closure Detection in Urban Environments. *International Journal of Advanced Robotic Systems*, 13(2), doi:10.5772/62755.

Wang, C. and Hsu, P.-H. (2007). Building detection and structure line extraction from airborne lidar data. *Journal of Photogrammetry and Remote Sensing*, 12(4):365–379.

Wiśniewski, Z. (2022). Total Msplit estimation. *Journal of Geodesy*, 96(10):82, doi:10.1007/s00190-022-01668-z.

Wiśniewski, Z. (2009). Estimation of parameters in a split functional model of geodetic observations (Msplit estimation). *Journal of Geodesy*, 83(2):105–120, doi:10.1007/s00190-008-0241-x.

Wiśniewski, Z. (2010). Msplit(q) estimation: estimation of parameters in a multi split functional model of geodetic observations. *Journal of Geodesy*, 84(6):355–372, doi:10.1007/s00190-010-0373-7.

Wyszkowska, P. and Duchnowski, R. (2019). Msplit Estimation Based on L 1 Norm Condition. *Journal of Surveying Engineering*, 145(3), doi:10.1061/(asce)su.1943-5428.0000286.

Wyszkowska, P. and Duchnowski, R. (2020). Iterative Process of Msplit(q) Estimation. *Journal of Surveying Engineering*, 146(3), doi:10.1061/(asce)su.1943-5428.0000318.

Wyszkowska, P. and Duchnowski, R. (2022). Processing TLS heterogeneous data by applying robust Msplit estimation. *Measurement*, 197:111298, doi:10.1016/j.measurement.2022.111298.

Wyszkowska, P. and Duchnowski, R. (2024a). Locally robust Msplit estimation. *Journal of Applied Geodesy*, 19(2):145–158, doi:10.1515/jag-2024-0023.

Wyszkowska, P. and Duchnowski, R. (2024b). Msplit Estimation with Local or Global Robustness Against Outliers—Applications and Limitations in LiDAR Data Processing. *Remote Sensing*, 16(23):4512, doi:10.3390/rs16234512.

Wyszkowska, P. and Duchnowski, R. (2025). Sliding window algorithm applied to Msplit estimation for seasonal change detection from LiDAR data. Technical report.

Wyszkowska, P., Duchnowski, R., and Dumalski, A. (2020). Determination of Terrain Profile from TLS Data by Applying Msplit Estimation. *Remote Sensing*, 13(1):31, doi:10.3390/rs13010031.

Yang, H., Omidalizarandi, M., Xu, X., and Neumann, I. (2017). Terrestrial laser scanning technology for deformation monitoring and surface modeling of arch structures. *Composite Structures*, 169:173–179, doi:10.1016/j.compstruct.2016.10.095.

Zhang, X., Wenjun, C., Zhang, X., Zheng, Y., Zhang, B., Shaoming, W., Jiandong, Y., and Guozhen, S. (2023). The Deformation Analysis of the 3D Alignment Control Network Based on the Multiple Congruence Models. *Journal of Geodesy & Geoinformation Science*, 6(2):21–31, doi:10.11947/j.JGGS.2023.0203.

Zhao, R., Pang, M., Liu, C., and Zhang, Y. (2019). Robust Normal Estimation for 3D LiDAR Point Clouds in Urban Environments. *Sensors*, 19(5):1248, doi:10.3390/s19051248.

Zienkiewicz, M. H. (2015). Determination of vertical indicators of ground deformation in the Old and Main City of Gdansk area by applying unconventional method of robust estimation. *Acta Geodynamica et Geomaterialia*, page 249–257, doi:10.13168/agg.2015.0024.

Zienkiewicz, M. H. and Dąbrowski, P. S. (2023). Matrix strengthening the identification of observations with split functional models in the squared Msplit(q) estimation process. *Measurement*, 217:112950, doi:10.1016/j.measurement.2023.112950.