Rigorous and extensive accuracy assestment of automatically classified LiDAR data: a case study in the city of Milan, Italy

Davide Lodigiani[1], Vittorio Marco Casella[1]

[1] Department of Civil Engineering and Architecture, University of Pavia - 27100 Pavia, Italy
(davide.lodigiani, vittorio.casella)@unipv.it

Keywords: Airborne LiDAR, Point Cloud Classification, ALS, Urban Areas, Airborne System, LiDAR.

Abstract

LiDAR data filtering has been an active research area for nearly thirty years and continues to present significant challenges due to the increasing density of acquired LiDAR data. This study analysed aerial LiDAR data from Milan, characterised by a density of 20-30 pts/m^2. Initiated in the summer of 2022, the survey included aerial and terrestrial surveys. Aerial LiDAR data was captured at a minimum of 20 pts/m^2 using the sensor Leica CityMapper-2S, while mobile data were acquired using Cyclomedia's MMS system across 2555 km of roads. The ALS dataset includes Milan's provincial territory and features unclassified and automatically classified point clouds in an industrial environment. Two areas, San Siro and Città Studi, were selected to create a ground truth without automated methods. We created a comprehensive ground truth dataset to validate our filtering method through valid and well-known algorithms like TerraSolid, and that obtained in an industrial environment where expert users applied these algorithms under time constraints. Our classification achieved 95.8% accuracy in the San Siro area and 94.6% in the Città Studi district, while the classification of the industrial environment obtained 93.7% and 88.9%, respectively. In the future, we intend to refine parameters to improve automatic classification accuracy and extend the process to other areas in Milan, integrating deep learning algorithms within the MATLAB environment.

1. Introduction

LiDAR data filtering has been a very active area of research for over 30 years and continues to be an extremely dynamic and constantly evolving environment in which the scientific community is concentrated and which remains an area of study of great relevance and interest. Technological advances have helped improve the quality and precision of data collected through LiDAR technology; infact, as technology advances, ALS LiDAR sensors are proposed to enable the acquisition of point clouds with increasingly higher densities. Consequently, new filtering methods for LiDAR data have been developed to handle the increasing density of points. As current effective filtering methods for low-density point clouds may not produce the same results with higher-density data. Research on LiDAR data filtering techniques is complex and ongoing. As LiDAR technology and its applications progress, it is essential to continually refine filtering methods to align with the accuracy and precision of new ALS LiDAR sensors. In addition, attention must be paid to the limited availability of ALS LiDAR datasets with complete and rigorous ground truths for classification. Among the datasets currently available and in use, we can find, for instance, the one related to the city of Vaihingen (DE) (Chakraborty and Dey, 2024, Özdemir et al., 2021, Feng and Guo, 2021). However, this dataset is characterised by a relatively low LiDAR data density, ranging from 4 pts/m^2 to 8 pts/m^2, and a limited extent.
Today, technology has evolved, and ALS LiDAR sensors can acquire cities with a point density much higher than the Vaihingen dataset. Nowadays, cities are acquired with sensors that can acquire points with densities ranging between 20 pts/m^2 and 40 pts/m^2. With this technological evolution of ALS LiDAR sensors, it is essential to have a test site that can cope with this acquisition method conceived today. Modern acquisition techniques offer increasingly greater spatial resolutions, generating large volumes of data with significantly higher point densities than previous systems. Filtering techniques must take this into account, with data that present a complexity given by the greater density that requires more advanced classification methods, provided by the greater amount of information associated with each point, and this translates into higher levels of accuracy and detail in the classification processes. In this regard, it becomes very interesting to have a rigorous LiDAR dataset that can adapt to the challenge of the acquisition method of these new technologies. Creating a LiDAR dataset, which includes different types of terrain, vegetation and infrastructure, will allow us to train machine learning models on this data type and enable us to be a solid base for the calibration and validation of classification algorithms. In this way, fully exploiting advanced acquisition technologies will be possible. This dataset will allow us to be a test bed for existing classification methodologies, allowing us to understand if they are robust enough to handle the complexity of modern data. As the density of points increases, classification techniques must recognise the various types of objects, such as terrain, buildings, vegetation and other infrastructure.
In the summer of 2022, the Municipality of Milan initiated an important project to survey the entire provincial area and create a digital twin of the city; this project included the acquisition of both aerial and ground data. The aerial surveys were carried out by CGR S.p.A. using the Leica CityMapper-2S hybrid sensor, which combines a 2 MHz LiDAR sensor with simultaneous nadir and oblique imagery The nadir images were acquired in four bands (red, green, blue and NIR), while the oblique images in three bands (red, green and blue). The nadir images have a resolution of 5 cm, while the LiDAR data were collected with a density of at least 20 pts/m^2. In parallel, mobile mapping was performed with Cyclomedia's MMS system, which collected 360° panoramic images and point clouds along the roads, covering 2555 km.
CGR S.p.A. managed and processed the aerial LiDAR data, par-

ticularly on data calibration, integrating radiometric information within the LiDAR flight swaths. In parallel, quality analyses of the LiDAR data were conducted to ensure the consistency and reliability of the collected measurements. Starting with the flight strips, which are initially segmented into tiles measuring 500x100 metres, a tiling process was carried out to produce tiles of 500x500 metres. Subsequently, an automatic classification algorithm was applied to the point cloud generated from the tiling process to classify the points into vegetation, terrain, water, bridges or viaducts, low points, power lines, and buildings.
The Municipality of Milan has authorized the Laboratory of Geomatics to use around 50 TB of survey data for research. The Laboratory of Geomatics has access to the following products:

- True Orthophoto: orthorectified aerial images, both RGB and CIR (Color InfraRed), with a resolution of 5 cm.
- Aerial nadir images with a 5 cm resolution and aerial oblique images.
- Raw aerial LiDAR point clouds: unclassified flight swaths, which contain radiometric information, are divided into tiles measuring 500x1000 m each.
- Digital Terrain Models (DTM) and Digital Surface Models (DSM) with a resolution of 50 cm.
- Point clouds from mobile terrestrial mapping (MMS).
- Aerial point clouds classified and divided into 500x500m square tiles.

For further information regarding the survey of the province of Milan, please consult the articles (Franzini et al., 2023, Franzini et al., 2024).
We are currently focusing on data obtained from aerial surveys, particularly LiDAR point clouds. In this regard, we have built a ground truth by manually classifying the points in two areas identified in Milan that have different characteristics. The first area is characterised by large green spaces; in contrast, a concentrated presence of buildings characterises the other.In addition to creating the ground truth, our goal was to use these rigorous datasets to validate the filtering performed by consolidated algorithms, which are considered valid and widely used, such as the one used in the commercial software TerraScan.The validation process involved two distinct methodologies: first, an industrial filtering performed by a company and, secondly, a filtering from scratch performed by us. Industrial filtering offers the significant advantage of being controlled by expert users who can manage the data efficiently. However, this method is often constrained by time, which can limit the attention to detail required for accuracy. On the contrary, the filtering we developed internally was implemented using the consolidated commercial software TerraScan, and the filtering has the characteristic of being built from scratch with particular attention to details typical of research.
In this article, the ground truth created for two study areas identified in the city of Milan will be illustrated. The process of creating the ground truth will be presented, describing in detail the methodologies and criteria used to obtain a rigorous dataset. Consequently, the results obtained from the ground truth created will be analysed in detail, focusing on the accuracy of the point cloud classification. Next, we will compare our classification with the filtering created in an industrial environment.

2. Materials

2.1 Point Clouds

The LiDAR dataset available to us includes the entire provincial territory of Milan, which covers an area of 1,776 km². The survey, carried out by the company CGR S.p.A., was performed with the Leica CityMapper-2S LiDAR ALS sensor, which can simultaneously acquire both LiDAR data, with a capacity of up to 2 million points per second and nadir and oblique images. The aircraft flew over the entire province at an average altitude of 1500 m AGL, providing point clouds with a density of at least 20 pts/m^2 and a nadir and oblique image resolution of 5 cm.
The available data also includes point clouds that have already been automatically classified in an industrial environment, covering the boundaries of Milan and the municipality of San Colombano al Lambro. These classified point clouds are organised in 500x500 metre tiles and saved in LAS v1.4 format. In addition to the classified tiles, raw LiDAR point clouds are also available, and each flight swath is subdivided into tiles of 500x1000 m.
The first area selected consists of 20 tiles measuring 500x500 m, while the second consists of 12 tiles of the same size. The point clouds contain radiometric information: each point is associated with the four spectral bands (red, green, blue, and NIR).

2.2 Study areas

The study areas are situated in the city of Milan, as shown in Figure 1. Milan, which is located at 45.46° N and 9.19° E, is the capital of the Lombardy region in northern Italy. Covering an area of approximately 182 km^2, Milan has a varied landscape, including rural, residential and high-density urban areas. On the outskirts, there are ample green spaces mixed with buildings, while the centre features densely built areas. This work evaluates the automatic classification of aerial point clouds related to two study areas: the San Siro area (peripheral area) and the Città Studi district (more central area). These two areas have different urban fabrics, allowing a rigorous evaluation of the classification process adopted.
The San Siro area, Figure 2, located in the north-western part of Milan, is one of the peripheral areas of the city, characterised by large green spaces and scattered buildings. This study area, with a total extent of 5 km^2, includes the districts of QT8, Lampugnano, San Siro, ex Fiera, San Leonardo, Portello, and Ghisolfa. The buildings in this area are primarily residential, with an average of four to five floors above ground; however, some buildings reach 25 floors.
The second study area, Figure 3, Città Studi, is located in the northeastern part of the city, closer to the city centre. The area covers 3 km^2 and is characterized by a prevalence of buildings compared to the San Siro area. This area is an academic and research centre, home to the university buildings and scientific institutes that give it its name. The area is a perfect example of architectural heterogeneity, with modern buildings coexisting with historic ones.

3. Methodology

The point cloud is classified using one of the programs from the TerraSolid suite. The programs run on CAD platforms (Bentley or Spatix) and consist of five modules: TerraScan, TerraMatch, TerraModeler, TerraPhoto and TerraStereo. The module

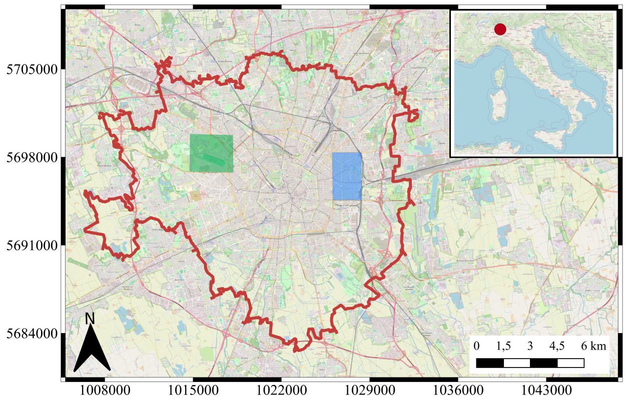

Figure 1. Map showing the two study areas and the Milan city border (in red): San Siro (green rectangle) and Città Studi (blue rectangle). [Map layer source: OSM, "Open Street Map", QGIS. Milan boundary source: Municipality of Milan Geoportal (SIT - Territorial Information System)].

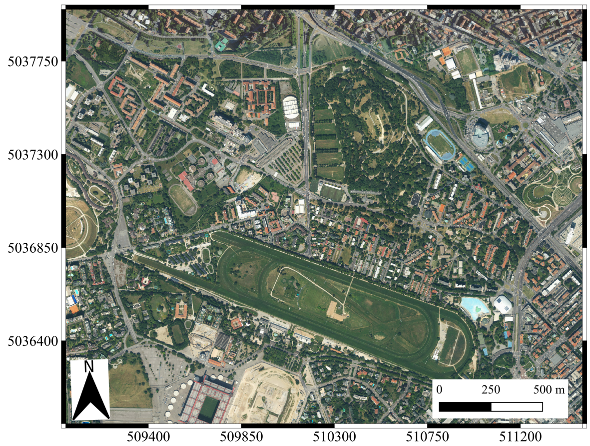

Figure 2. Orthophotos of the San Siro area taken during the flight in the summer of 2022. The image shows the urban fabric of the area. It is characterised by a prevalence of residential buildings with extensive green spaces (La Maura racecourse, San Siro Hill).

Figure 3. Orthophotos of the Cità Studi area obtained from 2022 flight. The image shows the urban fabric of the area, which is characterised by a high density of buildings, with a mix of historic and modern one. The area is characterised by a small number of green areas.

used in this work is TerraScan, a program for viewing and processing LiDAR data. This software includes LiDAR filtering, DTM/DSM modeling, point cloud classification, vectorization and several other functions. One major benefit of this software is its capability to create macros, which execute a series of operations in order, facilitating the automation of classification tasks (TerraSolid, 2025).
Figure 4 shows the classification process used in both study areas. The following paragraphs provide a detailed overview of the methodology adopted to develop the new filtering method we applied.

3.1 Ground Truth

Creating a ground truth is a complex and time-consuming process, especially for point clouds, as it requires a manual classification of each single point, assigning a label to each one. Unlike a 2D image, which provides a clear and simple context, a 3D context, such as a point cloud, presents a more complex challenge due to an unorganised nature of the 3D points. This implies that the operator cannot rely exclusively on the spatial information of the point cloud for an accurate classification, as it is insufficient to have the context of the points. For this reason, in creating the ground truth, the operator uses the images acquired simultaneously as the LiDAR data to have a visual context. This facilitates the correct interpretation and classification of each point that makes up the point cloud and allows for a classification to be identified that responds to the context of the point being examined, thereby reducing errors associated with the incorrect interpretation of the point cloud.
We selected six representative tiles from each study area to create the ground truth. As mentioned previously, this process is time-consuming; therefore, we opted not to classify every identified tile entirely; instead, we manually classified various objects, taking into account complex cases such as trees attached to buildings, complex roof shapes, cars beneath trees, and trucks and buses. We also considered simpler cases, including buildings with simple roof shapes, trees, and isolated cars, ensuring that our ground truth remained balanced between simple and complex instances. We analyzed the distribution of

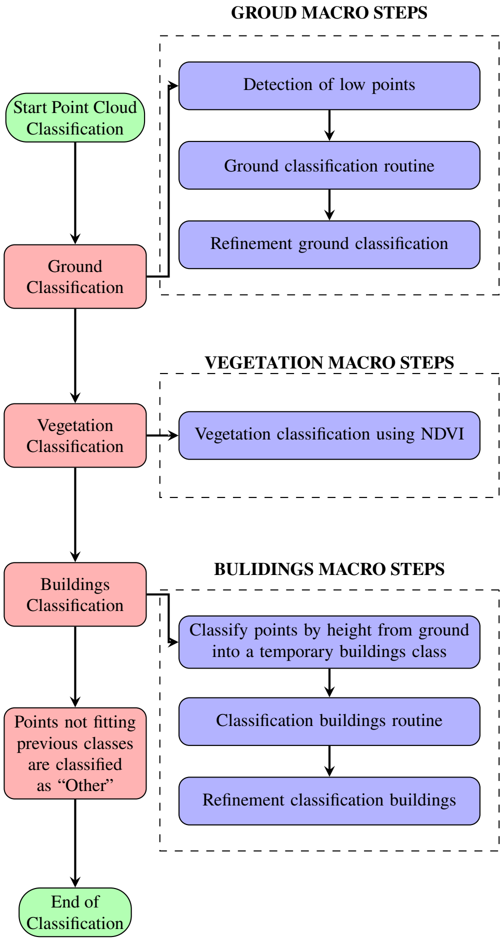

Figure 4. Flowchart showing the process of classifying aerial point clouds according to the macros implemented in TerraScan.

various classes in the two selected study areas using the classified tiles from the industrial environment. This allowed us to determine the percentage distribution of classes in both study areas, resulting in a ground truth that reflected reality. It is important to note that the six classified tiles as a whole correspond to the defined percentages, while each individual tile does not conform to this standard. The analysis indicated that the first study area, San Siro, is primarily characterised by ground and vegetation points. In contrast, the second study area, Città Studi, predominantly consists of building points. Table 1 displays the percentage distribution of the classes used to create the ground truth. In the San Siro area, 32 million out of 212 million points were classified (approximately 15%), while in the Città Studi district, around 14 million points out of 112 million were classified (about 13%).

The process we adopted in creating the ground truth is rigorous and based exclusively on an operator's manual classification. The classification process was carried out using the TerraScan software, which has numerous functions for manually classifying a point cloud. Typically, creating a ground truth for the point cloud begins with its automatic classification, followed by a manual refinement phase that serves to correct errors in the automatic algorithm. In this case, we relied exclusively on manual classification without using any automatic approach. This method has the disadvantage of being time-consuming; however, it is objective and neutral, and it avoids the influence of external factors without being influenced by the initial result of automatic classification. Figure 5 shows an example of one of the classified tiles.

The classes used for ground truth creation are as follows:

- Terrain: includes all points that represent the surface of the earth. This includes ground, pavements, roads and the bottom of watercourses.
- Vegetation: includes all points representing vegetation, including trees, bushes and grass at various heights. Vegetation on terraces or roofs is also included in this class.
- Buildings: represent roofs. This class does not include facades, which are part of the other class. This class also includes permanent roof objects such as chimneys, ventilation towers and antennas.
- Other: this class includes all points not covered by the previous classes. It includes unclassified or unidentifiable points, small structures that cannot be considered buildings, cars, poles, fences, building facades and all points considered to be noise.

Classes	**San Siro (%)**	**Città Studi (%)**
Other	10	22
Ground	45	28
Vegetation	33	21
Buildings	12	29

Table 1. Percentage of points in each class that serve as the basis for creating the ground truth for the two study areas.

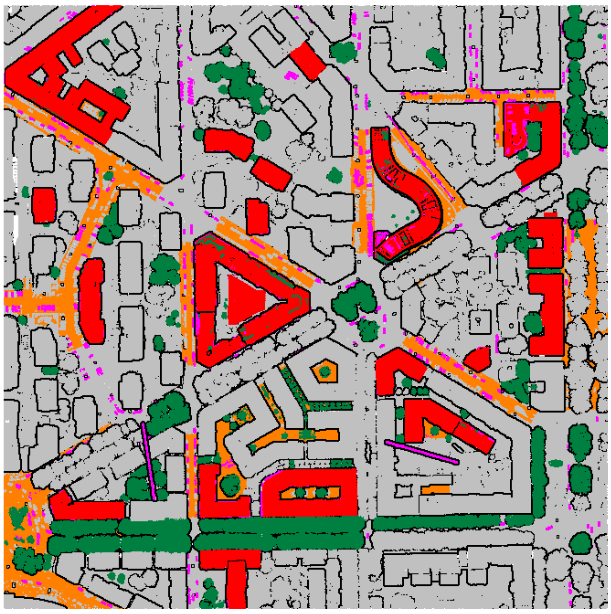

Figure 5. Example of the ground truth of the point cloud in the Città Studi district. This image represents one of the 6 classified tiles, illustrating the result of classifying the points according to the classes of interest. Buildings are represented in red, the ground in orange, the class called Other in magenta, vegetation in green and unclassified points in grey.

3.2 Ground Classification

Ground classification is the first step in point cloud classification. It is a fundamental step that creates the basis for many

applications ranging from DTM creation (Petschko et al., 2016) to 3D building vectorization (Albano, 2019).
Several scientific papers discuss the ground classification algorithm implemented in the TerraScan software (Brovelli and Lucca, 2012, Lin and Mills, 2009). The ground filtering algorithm implemented in TerraScan software software is based on the algorithm created by Peter Axellson (Axelsson, 1999, Axelsson, 2000) using TIN (Triangular Irregular Networks). The classification algorithm initiates the process by identifying seed points based on a user-defined grid, which includes the maximum building size as a parameter. The method then create TINs based on the previously determined points. With each iteration, additional points that meet the specified criteria are added to the triangle’s surface. The process stops when no additional points are added to the surface. Specific parameters must be set in the ground classification routine, such as:

- Terrain angle: maximum slope of the terrain
- Iteration angle: maximum angle between points
- Maximum building size: length in the plan of the largest building
- Iteration distance: to avoid creating too large TIN triangles

In addition, iteration angles that are too large may include points that have nothing to do with the ground, such as vegetation. When starting the ground classification routine in TerraScan software, it is important to note that the parameters are preconfigured with default values. However, it is important to modify these parameters according to the topographical characteristics of the study area. The TerraScan software manual (TerraSolid, 2025) suggests value ranges for the parameters described as follows:

- Iteration distance: values between 0.5 m and 1.5 m.
- Terrain angle: values between 88° and 90° (if there are man-made objects), and sum values of 10° or 15° degrees for nude terrain with only vegetation.
- Iteration angle: small values close to 4° for flat terrain and values close to 10° for mountain terrain.

The macro we created for ground classification starts with identifying low points. These points are essential before the ground classification, as they allow the ground to be correctly classified and prevent the areas below the ground from being classified as ground. Then, as recommended in the TerraScan manual, the candidate points for ground are identified using the ”hard surface” function. This allows the step of identifying all the points that could be classified as ground, reducing the number of points the algorithm has to analyze and, therefore, reducing the classification time. The next step is classifying the ground based on the previous candidate’s points for ground. The points not classified as ground are reclassified to the undefined class once the ground has been identified. Finally, all points within 5 cm of the ground are classified as ground.
The table 2 shows the parameters used to classify the ground in the two study areas. The parameters are identical except for the maximum size of the building. This parameter is particularly relevant in the Città Studi study area: an incorrect setting can cause an incorrect ground classification. Small areas, like building courtyards, may not be recognised as ground, while structures close to the ground, such as garages, might be misclassified as ground.

Parametres	**San Siro**	**Città Studi**
Max building size (m)	125.0	50.0
Terrain angle (°)	88.0	88.0
Iteration angle (°)	12.0	12.0
Iteration distance (°)	1.0	1.0
Reduce iteration angle when edge length (m)	5.0	5.0

Table 2. Parametres chosen for the classification of the ground for the two study areas.

3.3 Vegetation Classification

In addition to the three R, G and B bands, the point cloud attributes also include the NIR (Near InfraRed) band. With four bands available, it is possible to use a range of indices to identify vegetation. The most common index used in literature to determine the presence of vegetation is the Normalized Difference Vegetation Index (NDVI).
The NDVI parameter is generally used to determine the health and vigour of vegetation. This index uses two bands: the red band and the NIR band. Healthy plants absorb most of the red light for photosynthesis while they reflect the light in the Near InfraRed. NDVI uses this difference in reflectance to evaluate the state of vegetation and its presence. NDVI values vary from -1 to +1; a value close to -1 indicates no vegetation presence, while values close to +1 indicate dense and lush vegetation. The NDVI (Huang et al., 2021) formula is:

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

In both cases, there is a significant peak in NDVI around 0, indicating that most of the points in the two study areas do not represent vegetation. Analysing the interval between 0 and 1, it can be observed that the presence of points associated with vegetation is more significant in the San Siro area than in Città Studi.
For both study areas, a minimum threshold of 0.15 was chosen for the NDVI index, in line with what is reported in the literature for the identification of vegetation (Hashim et al., 2019), while the maximum value was set at 1.

3.4 Buildings Classification

Building classification is a fundamental step in creating 3D vector models of buildings. In this process, the facades of the buildings are not classified since this information is not considered by the software currently in use. The TerraScan software utilised for vectorisation generates vector models of buildings with a Level of Detail (LOD) of 2. Even if the facades were classified, the algorithm would not account for their position on the roof, as it first generates the slopes and then calculates the walls from the perimeter of the roof.
The macro developed for building classification starts by classifying all points into a temporary class, where all points in that class are potential candidates for being a building. This is achieved by classifying all points higher than 2m above ground level. After this step, any points not identified as buildings are returned to a default class. To include fixed objects on the roof, such as chimneys, the “group by best match” function is used to classify these objects into the building class as well. The building classification algorithm implemented in the TerraScan software is based on the algorithm developed by Peter Axellson.

The algorithm utilises information about the height of points from the ground: roofs are typically flat surfaces, and thus, changes in height along the scan lines are observed. If the elevation changes little along a scan line (i.e., the elevation changes do not vary significantly), it is likely to be a roof; on the other hand, considerable variations indicate vegetation or changes in the direction of the roof. The model searches for flat surfaces and identifies roofs using a simplified method that examines elevation changes along the scan lines. When the second derivative of elevation along the scan line is zero, it indicates uniform elevation changes on a flat surface like a building's roof. Conversely, a non-zero second derivative denotes a change in slope, reflecting alterations in the roof's direction or uneven surfaces, potentially due to vegetation or a roof ridge. The method then employs the second derivative to detect these changes and to distinguish between flat surfaces (indicating buildings) and rough surfaces (indicating vegetation or other non-flat objects). Table 3 shows the values of the parameters used to classify buildings in the two study areas.

Parametres	**San Siro**	**Città Studi**
Accept using	Normal rules	Relaxed rules
Min. size building (m^2)	50.0	25.0
Z tolerance (m)	0.25	0.25

Table 3. Parameters chosen for the classification of the buildings for the two study areas.

4. Results

The LiDAR point cloud classification results, are illustrated in Figure 6 and Figure 7, which respectively show a portion of the ground truth and a comparison between our automatic classification and that obtained in an industrial environment. Confusion matrices and Cohen's kappa parameters are reported for each approach. Figure 6 shows a building in the study area of San Siro and a small part of the surrounding area.. In contrast, Figure 7, which identifies a portion of the Città Studi district, shows a more extensive reference area with a greater number of buildings and other objects. The classification considered four classes: other, ground, buildings, and vegetation. In analysing the first study area, the San Siro area, the automatic classification method in an industrial environment highlights high user accuracies for the ground and vegetation classes, with a reduction in accuracy for the buildings and other classes. The producer's accuracy indicates a good ability to identify the ground and vegetation classes, with strong performance for the building class and difficulties for the class called other. Our classification method demonstrates high accuracy for the ground, vegetation, and buildings classes, while the other class has low accuracy. Regarding producer accuracy, our approach improves compared to the first method for the buildings class, while the other class continues to face problems. In this first study area, we can conclude that our adopted method is more accurate for the building class. However, both methods indicate that the classification of the other class is critical for both methods.

Moving on to the second study area, the Città Studi district, the classification method in an industrial environment highlights high user accuracy for the ground and buildings classes, followed by the vegetation class. In contrast, the other class has a low accuracy value. As for the producer accuracy, good recognition capabilities of the ground class are highlighted, followed by vegetation and buildings; however, in this case, the producer accuracy of the other class is very high, but the user accuracy is low (less than half), which suggests that the algorithm struggles to identify the points that are part of the other class. Our approach shows high user accuracy for the ground, buildings, and vegetation classes, while the other class confirms lower accuracy. The user accuracy is high for the vegetation and ground classes, followed by buildings, while the other class is characterised by lower accuracy (about 78%).

A significant challenge in classifying aerial point clouds is distinguishing structures such as bridges and viaducts from buildings or terrain. The misinterpretation is often due to the similarity in appearance between elevated viaducts and buildings due to their flat surfaces, leading to misclassification by algorithms. Currently, TerraScan software does not have a fully automated method for classifying these structures using point clouds alone; existing techniques rely on complementary data, such as polygons that indicate locations and boundaries of bridges.

Another challenge is accurately distinguishing vegetation next to buildings. We are exploring a strategy that starts with filtering points with a single return, typically corresponding to ground or roof surfaces. This initial step may help reduce misclassification risks, though it does not provide information on objects located on roofs. Additionally, we are considering classifying points according to specific planes to isolate vegetation for better classification.

We achieve a high accuracy for buildings; however, there are cases where objects like trucks are misidentified as buildings due to their horizontal surfaces. To avoid this problem, increasing the area parameter within the building classification routine may be a solution; however, this could result in the loss of some building details. Furthermore, the use of polygons to identify buildings is not possible as the available information is outdated and typically does not align with the building perimeters in LiDAR data. Our focus is on refining classification by utilizing TerraScan's grouping function to minimize misclassification errors; we aim to adjust parameters so that vehicles are identified in "other" class rather than buildings one.

Both methods show difficulties in the class "other". This difficulty is probably related to the composition of this class, as it includes poles, cars, building facades, and in general, all the points that do not belong to the other defined classes. The confusion matrices show high accuracies for the ground and buildings classes, but some confusion is noted in the other and vegetation classes. To improve the overall accuracy, a possible approach could be to refine the separation between the classes, particularly the vegetation class, which shows general confusion with the other classes and buildings.

5. Conclusion

Creating ground truth is a time-consuming process, but it enables us to evaluate the effectiveness of an automatic classification method objectively. In this article, we compare two classification methodologies: one used in an industrial setting and the other in a research context. To accomplish this, we identified two study areas in Milan, each with distinct characteristics; this allowed for an adequate evaluation of the classification process.

The comparison between the industrial method and our approach reveals some differences: our methodology generally demonstrates higher global accuracy. In the first study area, we achieved a global accuracy of 95.8%, compared to 93.7% in the industrial setting. In the second study area, slightly lower accuracy rates were recorded for both methods, precisely 94.6%

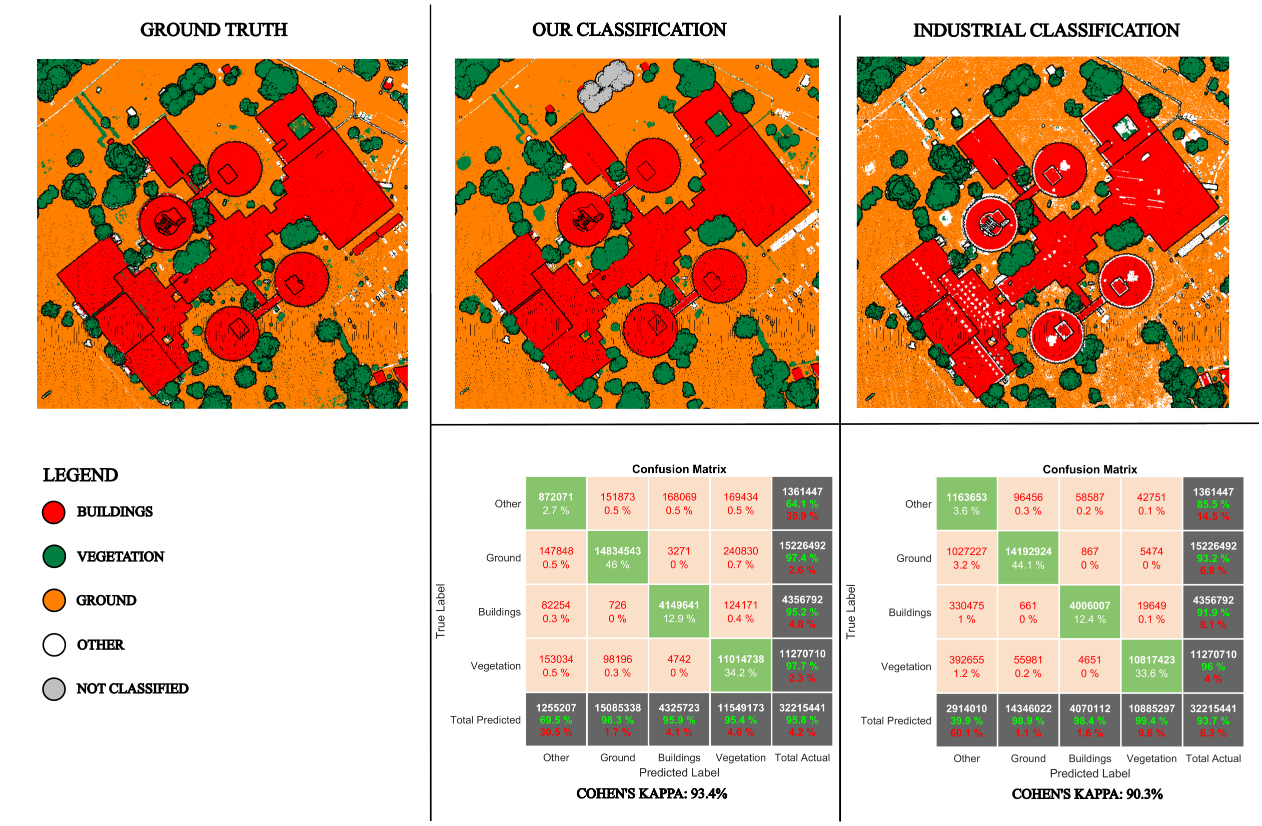

Figure 6. The ground truth, compared with the two classification techniques applied in the San Siro study area, along with the corresponding confusion matrix, illustrates the performance of these classification methods.

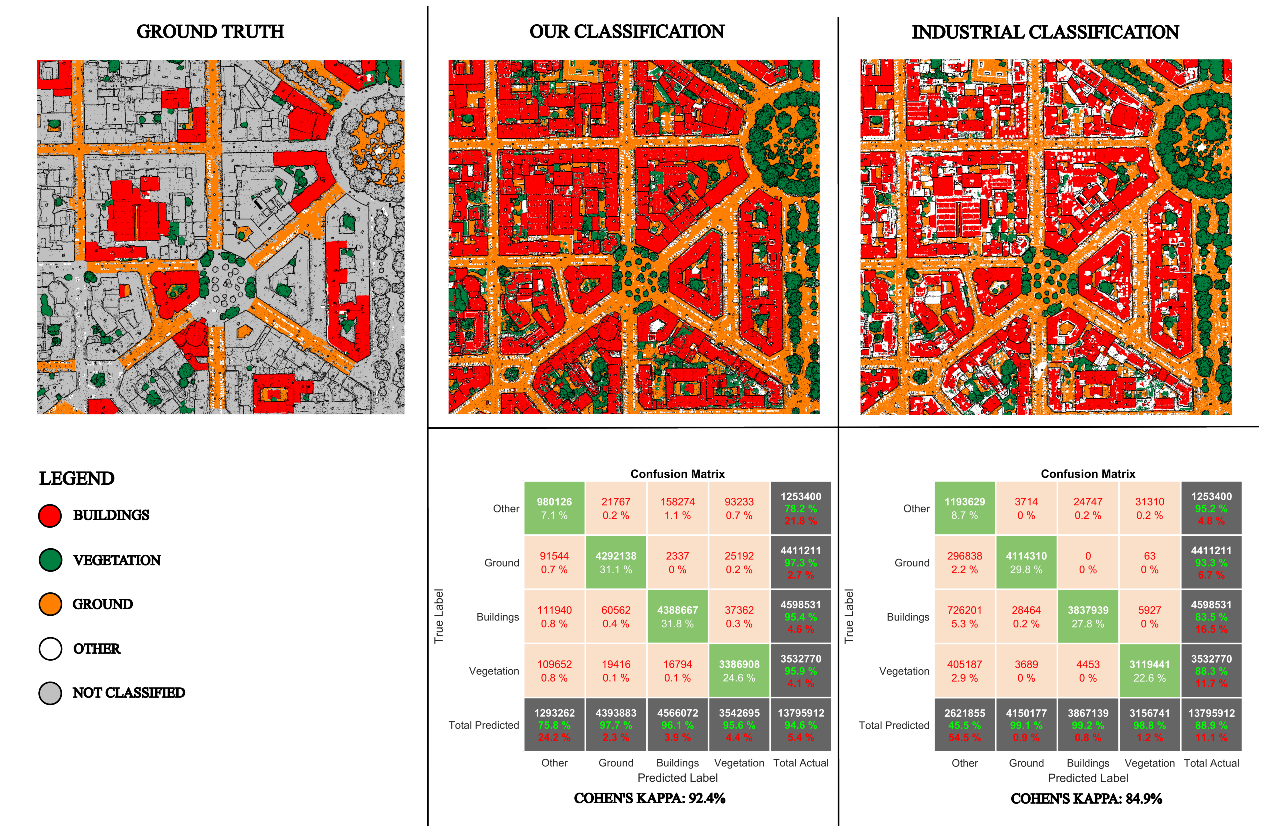

Figure 7. The ground truth, compared with the two classification techniques applied in the Città Studi district, along with the corresponding confusion matrix, illustrates the performance of these classification methods.

for our classification and 89% for the industrial classification. This variance between the two study areas can be attributed to environmental complexity, particularly in the Città Studi district, which features adjacent buildings and complex shapes. The TerraScan software algorithm emphasises the generalisation capability that distinguishes the various objects in the two study areas, supported by confusion matrices that illustrate the robustness of the classification process.

Future developments will certainly involve further refinement of the parameters to improve the accuracy of automatic classification. Concurrently, we will continue creating the ground truth by classifying the remaining tiles for the two study areas using the same methodology adopted in this study. Additionally, we will extend this classification process to other areas of Milan, allowing us to assess the algorithm's performance in various scenarios, such as the city centre with its complex urban fabric. Another future goal is to utilise deep learning algorithms to classify the aerial point cloud automatically. We have a rigorous, manually classified ground truth, currently including six tiles for each study area, which we plan to extend by classifying the remaining tiles in the two study areas using the same methods employed to classify the six tiles. This dataset not only facilitates the validation of the automatic classification process with commercial software but also serves to train deep learning algorithms. In particular, the MATLAB environment will be utilised to train these algorithms: MATLAB already includes a Deep Learning module which, specifically for point clouds, features pre-trained neural networks such as RandLA-Net (Hu et al., 2020) and PointNet++ (Qi et al., 2017) on the DALES dataset. With an existing pre-trained neural network, a viable approach involves applying the transfer learning technique, which allows us to leverage a network that has already been trained and adapt it to our dataset without necessitating a complete training process from scratch. The results obtained from deep learning will then be compared with established algorithms like those available in the commercial software currently used.

In conclusion, our classification process has shown encouraging results, with accuracies exceeding 94% for both study areas. Analysing two study areas with distinct urban characteristics has allowed us to validate the TerraScan software algorithm rigorously, ensuring satisfactory results across various scenarios.

Acknowledgements

The authors thank the Municipality of Milan for sharing data used in this study for research purposes. Special thanks to Arch. Bruno Monti for his valuable support and collaboration during this process.

This paper was prepared within the Doctoral Programme Cycle XXXIX, funded by the Ministry of Universities and Research (MUR) through the National Recovery and Resilience Plan (PNRR).

The software used in this study was bought through the NODES project, funded by the European Union under the NextGenerationEU programme. Specifically, the NODES project is integral to Mission 4, Component 1.5, with the code ECS00000036 and the CUP F17G22000190007.

References

Albano, R., 2019. Investigation on roof segmentation for 3D building reconstruction from aerial LIDAR point clouds. *Applied Sciences (Switzerland)*, 9.

Axelsson, P., 1999. Processing of laser scanner data - Algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54.

Axelsson, P., 2000. DEM Generation from Laser Scanner Data Using adaptive TIN Models. *International Archives of Photogrammetry and Remote Sensing*, 23.

Brovelli, M. A., Lucca, S., 2012. Comparison of GRASS-LiDAR modules-TerraScan with respect to vegetation filtering. *Applied Geomatics*, 4.

Chakraborty, D., Dey, E., 2024. Segmentation of LiDAR point cloud data in urban areas using adaptive neighborhood selection technique. *PLOS ONE*, 19.

Feng, C.-C., Guo, Z., 2021. A Hierarchical Approach for Point Cloud Classification With 3D Contextual Features. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 5036-5048.

Franzini, M., Casella, V. M., Monti, B., 2023. Assessment of Leica CityMapper-2 LiDAR Data within Milan's Digital Twin Project. *Remote Sensing*, 15.

Franzini, M., Casella, V., Monti, B., 2024. The digital twin of the metropolitan area of milan: Quality assessment of aerial and terrestrial data. E. Borgogno Mondino, P. Zamperlin (eds), *Geomatics for Environmental Monitoring: From Data to Services*, Springer Nature Switzerland, 152–171.

Hashim, H., Latif, Z. A., Adnan, N. A., 2019. Urban vegetation classification with ndvi threshold value method with very high resolution (vhr) pleiades imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42.

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A., 2020. RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Huang, S., Tang, L., Hupy, J. P., Wang, Y., Shao, G., 2021. A commentary review on the use of normalized difference vegetation index (ndvi) in the era of popular remote sensing.

Lin, Y.-C., Mills, J. P., 2009. Integration Of Full-Waveform Information Into The Airborne Laser Scanning Data Filtering Process. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII.

Petschko, H., Bell, R., Glade, T., 2016. Effectiveness of visually analyzing LiDAR DTM derivatives for earth and debris slide inventory mapping for statistical susceptibility modeling. *Landslides*, 13.

Qi, C. R., Yi, L., Su, H., Guibas, L. J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *arXiv preprint arXiv:1706.02413*.

TerraSolid, 2025. Terrascan user guide 64 bit terrascan. `https://www.terrasolid.com/guides/tscan/index.html`. Last access: 2025-01-21.

Özdemir, E., Remondino, F., Golkar, A., 2021. An Efficient and General Framework for Aerial Point Cloud Classification in Urban Scenarios. *Remote Sensing*, 13, 1985.