Politecnico di Torino

Master's Degree Thesis

From Aerial Photogrammetry to Per-Building Rooftop PV Suitability for the Historic Centre of Turin

Candidate: Pouya Rezaei

Supervisor: Prof. Antonia Spanò

Co-supervisors: Prof. Giacomo Patrucco, Prof. Guglielmina Mutani

Academic Year 2025/2026

Contents

Abstract

This thesis investigates the realistic potential for rooftop photovoltaic (PV) capacity that could be installed on the buildings of a study area within the historic centre of Turin, and evaluates the reliability of the standard Italian method for estimating usable roof area at the level of the individual building. In Italian practice, the PV-usable area is typically estimated by multiplying the building footprint by a fixed availability coefficient of about 0.35 of the footprint, a simplified approach commonly used for city-scale estimates. While this approach is convenient for city-scale assessments, it has not been validated at the building scale — the scale at which incentive allocation and heritage decisions are actually made.

The analysis begins with the 2018 aerial survey of the city, from which a photogrammetric point cloud of about 654 million points was generated. This cloud was classified into ground, building and tree categories with a deep-learning segmentation model that achieved an overall accuracy of about 98.5%, and the buildings were reconstructed in three dimensions at Level of Detail (LOD) 2.1. On the reconstructed roofs, a classifier identified the clean, usable surfaces by excluding chimneys, dormers and other obstructions, and this usable area was combined with solar irradiation computed with the r.sun model. The analysis was structured as a four-layer cascade: gross roof area, available area, suitable (obstruction-free) area, and the resulting solar energy.

A comparison between the suitable area and the 0.35 coefficient shows that the coefficient over-estimates the truly usable area by about 44% across the historic centre, and that at the individual-building level it is unreliable, ranging from well below to more than twice the measured value; this apparent agreement at the aggregate scale conceals large per-building discrepancies. The analysis was then extended to hourly self-consumption for the residential buildings, using census family counts and standard demand profiles. The study area proves to be PV-undersized: for a typical building most of the production is consumed on site (median self-consumption about 70%), yet rooftop PV can supply only about one quarter to one third of the annual electricity demand. The principal contribution of this work is the disaggregation of a single bundled coefficient into distinct, measurable steps, applied building by building within an Italian historic centre.

Keywords: rooftop photovoltaic; urban building energy modelling; LOD2.1; point-cloud segmentation; self-consumption; historic centre; Turin.

Chapter 1

Introduction

Cities play a central role in the energy transition, and the rooftops of their buildings are one of the largest surfaces still available for renewable energy generation. Rooftop photovoltaic (PV) is therefore one of the main options for producing clean electricity within the urban fabric. Historic centres, however, present a special and difficult case: they are densely built, protected for their heritage value, and their roofs are old, complex in geometry and often oriented in irregular directions. These same features make them important — they hold a large share of a city's building stock and energy demand — and harder to assess than a modern district, where roofs are larger, simpler and better oriented.

Assessing rooftop solar potential across a whole district is not possible without a three-dimensional model of the buildings. In recent years, 3D city models have become an important instrument for governing and planning cities: public administrations use them to monitor the built environment, to run large-scale analyses, and to forecast quantities such as energy demand, solar potential and environmental risk (Biljecki et al., 2015a), and recent studies continue to use them for high-resolution, large-scale rooftop PV assessment (Winkler et al., 2025). A 3D city model can be produced in several ways — from airborne LiDAR, from aerial or satellite photogrammetry, or by extruding existing cadastral maps — and each source gives a different balance of cost, coverage and detail. In this work the models are built from aerial photogrammetry, which is comparatively low-cost, widely available, and provides true colour together with dense geometric detail.

In Italy, the usable rooftop area for PV installations is typically estimated by multiplying the building footprint by a fixed availability coefficient of about 0.35. This coefficient-based approach was introduced for the Italian building stock by Bergamasco & Asinari (2011a, b) and is the one used in the urban-energy methodology developed by the research group of Prof. Guglielmina Mutani that this thesis follows, which takes the usable roof as a comparable fraction of the building (Mutani & Todeschi, 2021; Mutani et al., 2023). While this method is convenient and performs reasonably well at city scale, it has never been systematically validated building by building. Yet the single building is exactly the scale at which incentives are granted and heritage decisions are taken, so a method that is only reliable on average is not enough.

The aim of this thesis is to measure the usable rooftop PV area building by building in the historic centre of Turin, starting from aerial photogrammetry, to test whether the 0.35 coefficient holds at that scale, and to extend the analysis to hourly self-consumption. To reach this aim, the work has five objectives: (1) to classify the photogrammetric point cloud into ground, building and tree; (2) to reconstruct the buildings in three dimensions at level of detail LOD2.1; (3) to measure the suitable roof area of each building — the part that could actually host photovoltaic panels, once obstructions and fragments too small for a panel are removed — as a layered cascade; (4) to compare this measured area with the 0.35 coefficient; and (5) to model the hourly Self-Consumption Index (SCI) and Self-Sufficiency Index (SSI) of the residential buildings.

The study area is a delimited portion of the historic centre (*centro storico*) of Turin — not the entire centre. The work starts from the municipal aerial photogrammetric survey of the city, which produced a point cloud of about 654 million points. From these data, 1396 buildings were reconstructed out of 1510 footprints. A particular feature of the centre is that its roofs follow the orthogonal Roman street grid, which is rotated to run along the intercardinal directions, so most of them face intercardinal rather than due south.

The work follows a single chain of steps: aerial photogrammetry, deep-learning classification of the point cloud, reconstruction of the buildings at LOD2.1, a roof-suitability classifier combined with solar irradiation computed with the r.sun model, a four-layer usable-area cascade, and finally an hourly self-consumption analysis. Each step is described in detail in the chapter that deals with it.

The main finding is that the 0.35 coefficient over-estimates the truly usable roof area by about 44.5 % across the whole study area, and that building by building it is unreliable, with the ratio between the coefficient and the measured area ranging from about 0.70 to 2.42. The apparent agreement at city scale therefore hides large per-building errors. The self-consumption analysis shows that the centre is PV-undersized: for a typical building most of the production is used on site (median Self-Consumption Index around 70.5 %), but rooftop PV can supply only about a quarter to a third of the annual electricity demand (city-wide Self-Sufficiency Index around 27 %). The main contribution of this work is to break a single bundled coefficient into separate, measurable steps and to apply them building by building to an Italian historic centre. Most of the tools used are existing and open — for example the roof-segmentation method of Usta et al. (2025a) — and the original part is their integration into one pipeline, the classifiers trained on data labelled specifically for this thesis, the interactive 3D viewer built to explore the results, and the per-building comparison.

A further and distinctive element of this work is the *scale of the input data*. Existing city-scale assessments, such as Usta et al. (2025a, 2025b), work on a digital surface model of about 1 m resolution. This thesis instead starts from an aerial photogrammetric survey and builds a much finer, decimetre-level surface model — a ground sampling distance of about 7.57 cm and a roof analysis carried out at 5 cm. It is precisely this finer scale that motivated the photogrammetric, geomatics-based approach and the more detailed, two-step roof segmentation adopted here: the level of detail of the data is decisive for a building-by-building analysis, and the photogrammetric DSM makes it possible to resolve individual roof faces, dormers and obstructions that a 1 m model cannot capture. In this sense the work also connects to the built-heritage and 3D city-model modelling developed in the geomatics laboratory (Aboohamzeh, Avena & Spanò, 2024).

This work is built on open and freely available tools, and the code developed for it is released openly. This follows the FAIR principles, which state that research outputs should be Findable, Accessible, Interoperable and Reusable (Wilkinson et al., 2016). The full pipeline — the source code, the processing scripts for each step, and the tool configurations — is available in a public repository under an open licence[1]; the input data are described

[1] `https://github.com/rezaeipouya97/rooftop-pv-suitability-turin`

in detail but cannot be redistributed.

The thesis is organised as follows. Chapter 2 reviews the state of the art for the four parts of the work. Chapter 3 describes the aerial photogrammetry and the deep-learning classification of the point cloud. Chapter 4 covers the reconstruction of the buildings at LOD2.1. Chapter 5 measures the usable rooftop area building by building and compares it with the 0.35 coefficient. Chapter 6 extends the analysis to the hourly self-consumption of the residential buildings.

Chapter 2

State of the Art

2.1 Aerial photogrammetry and point-cloud classification

Photogrammetry is the science and process of extracting 3D information and precise measurements from photographs taken either from the air or, in close-range applications, from a terrestrial perspective. Nowadays digital photogrammetry integrates the Structure-from-Motion (SfM) technique — implemented in photogrammetric software and based on image-matching algorithms — to automate the generation of the 3D model. SfM reconstructs a 3D model from a set of overlapping images taken from different viewpoints, and it works in two stages. The first stage, called correspondence search, looks for the same scene across the input images and organises the matches into a scene graph. Building on this scene graph, the second stage, called incremental reconstruction, iteratively recovers the camera positions and the 3D structure (Schönberger & Frahm, 2016). Compared with LiDAR (Light Detection and Ranging), photogrammetry uses passive optical sensors and is cheaper and more widely available, and it provides true RGB (red, green, blue) colour (Chiabrando, Costamagna & Spanò, 2013); unlike LiDAR, however, photogrammetry depends on the illumination conditions at the moment of acquisition, and the chosen georeferencing strategy can introduce further uncertainties.

The photogrammetric points are produced from high-resolution images captured from above the buildings. Compared with LiDAR, this point cloud is less dense and noisier, but it is uniformly distributed and complete over the visible surface (Can et al., 2021). The raw cloud carries only the position (XYZ) and colour (RGB) of each point: it does not record which points are ground, which are buildings and which are trees. Before the cloud can be used, it therefore has to be classified into these classes.

Because a city-scale cloud contains hundreds of millions of points, the classification method has to be efficient. A standard dense convolutional network is very inefficient on this kind of data, because most of the 3D space is empty and the network still spends computation on that empty volume. Submanifold Sparse Convolutional Networks (SSCNs) are far better suited to large point clouds: they store and process only the occupied voxels and skip the empty ones, which makes the computation faster and much lighter (Graham et al., 2018). Sparse convolution has two main open-source implementations: MinkowskiEngine (Choy et al., 2019), the first generalised sparse-convolution library, and TorchSparse (Tang et al., 2022; Tang et al., 2023), which improves its training and inference efficiency. The networks are usually built in an encoder–decoder (U-Net) form (Ronneberger et al., 2015).

Most recent work on point-cloud semantic segmentation has been carried out on LiDAR data. Aerial photogrammetry is more cost-effective for city-scale acquisition, but it brings its own difficulties: the cloud describes only the visible surface, and it is noisier than a LiDAR scan. Benchmark results on photogrammetric urban datasets remain modest (Can et al., 2021), and the problem becomes harder still in a dense historic centre, where the roofs have complex geometry and architecture.

Ready-to-use classification tools, such as the built-in classifiers in standard photogrammetric software, are designed for cleaner or already-classified data and are validated mainly on simpler scenes. On a dense, noisy historic-centre cloud they do not reach the precision that the later steps of this work require. For this reason, instead of relying on a general-purpose model, a dedicated classifier was trained for the photogrammetric point cloud of the study area in the Turin historic centre.

2.2 Three-dimensional city models and building reconstruction

Nowadays, 3D city models have become the standard way of representing the urban environment and a key instrument for governing and planning cities: public administrations use them to monitor the built environment, to run large-scale analyses, and to forecast quantities such as building energy performance, solar potential and environmental risk (Biljecki et al., 2015a). The geometric detail of these models is described by the Level of Detail (LOD) concept introduced in the CityGML standard, which ranges from LOD0 (only the 2D building footprint) up to LOD4 (full detail, including the interior). To remove the ambiguity of this coarse scale, it was later refined into sub-levels, from LOD2.0 to LOD2.3, which state more precisely how accurately the building shape and the roof geometry are reconstructed (Biljecki et al., 2016), a scheme later revised in the CityGML 3.0 standard (Kutzner et al., 2020). Because estimating the energy production of a roof requires its exact shape, geometry and slope, a model at LOD2 is essential for this kind of analysis.

In recent years, the field of city modelling has moved towards automated reconstruction at large city and even national scale. These automated pipelines combine the point cloud with the building footprints and can reconstruct millions of buildings at different LODs at the same time. A clear example at national scale is the work of Peters et al. (2022), who developed a workflow that automatically reconstructs the 3D models of all the buildings in the Netherlands. This shows that reconstructing buildings from a point cloud and footprints has reached a level of maturity and has become an established and proven approach, ready to be used on large-scale city datasets.

Even so, applying these automated pipelines is not straightforward. Most of them, and their default parameters, were designed and calibrated for airborne LiDAR data, as is the case for the pipeline of Peters et al. (2022). Both LiDAR and photogrammetry capture the visible outer surface of the scene, and airborne LiDAR generally gives the cleaner, better-structured cloud with well-understood noise; the acquisition distance is, in either case, one of the crucial factors for the quality of the result. A photogrammetric cloud, by contrast, is noisier and carries more geometric noise (Can et al., 2021). For this reason the tool cannot simply be used out of the box with its default settings; the parameters have to be adjusted for the photogrammetric data.

Once the buildings are reconstructed, the result has to be stored in a standard format for data exchange; for this purpose CityJSON has been introduced as an efficient and standardised encoding of the CityGML model (Ledoux et al., 2019). Beyond the choice of output format, however, the more important issue is that the level of detail itself directly influences the result of the spatial analysis. Studies have shown that the geometric accuracy and detail of the 3D model propagate directly into the estimated solar irradiation of the roofs, so the level of detail of the model has a real effect on the computed energy potential (Biljecki et al., 2015b). This is exactly the question this study addresses: reconstructing a dense historic centre from photogrammetric data, and examining how much the level of detail changes the roof area that is usable for PV.

2.3 Estimating the usable rooftop area

With the 3D geometry of the roofs available, the next step is to calculate the area on which PV panels can actually be installed. Reaching the usable area from the raw roof geometry is not straightforward: the gross roof area is never equal to the area that can be used for PV. Several factors reduce it, such as physical obstructions (chimneys, dormers, skylights), orientation and slope, and shading from neighbouring buildings. For this reason, estimating the suitable area for photovoltaic systems in an urban environment is recognised as a research problem in its own right (Freitas et al., 2015).

In previous studies, the usable rooftop area has been estimated along a range of methods, from coarse to fine. The simplest approach multiplies the building footprint, or the gross roof area, by a constant coefficient. This is useful for a quick assessment at urban or national scale, but because it ignores the real architecture of the building, its accuracy is low. A second, more detailed approach is based on the digital surface model (DSM) in a GIS (Geographic Information System) environment, where the solar irradiation is computed for each pixel and the surfaces are filtered by a threshold on slope and aspect. On top of these, a third group of methods applies reduction, or usability, factors that lower the gross area to account for the physical obstructions and for the spacing needed to install and maintain the panels. Each of these approaches simplifies the complex geometry of the roofs to a degree that suits the scale of the study. The present work builds on the DSM-based approach and applies such reduction factors as part of its analysis.

For an accurate assessment at the building scale, a more specialised approach divides the roof into planar segments according to slope and aspect. Instead of treating the whole roof as a single unit, each homogeneous segment is evaluated with its own potential. This segmentation is an efficient and established way to handle the complex geometry of real roofs, and it was used, for example, by Usta et al. (2025a) to extract the usable area; more generally, the automated modelling of buildings and 3D city models from point clouds is an active line of work in the geomatics field (Aboohamzeh, Avena & Spanò, 2024). The present study follows the same segmentation method in the first step of its roof analysis.

Nevertheless, the coarser methods — a fixed footprint coefficient, or a simplified flat-roof model — are usually calibrated on modern cities or on buildings in the periphery. Applying a single coefficient to a dense historic centre, with its complex roof geometry, unusual orientations and large number of small obstructions, is therefore highly unreliable. Fine, per-face assessment of the usable area from high-resolution data is rarely applied to this kind of fabric. The aim of this study is to evaluate the usable roof area face by face and to provide it for a dense historic centre, where the previous general and coefficient-based methods cannot give a reliable estimate.

2.4 Self-consumption and urban building energy modelling

Estimating the annual producible energy alone does not show whether the energy is produced at the time it is needed. PV production reaches its peak in the middle of the day, while for residential use the electricity is mostly demanded in the evening (Luthander et al., 2015). Because of this temporal mismatch, evaluating the performance of a solar system has to shift from the quantitative question of “how much energy is produced?” to the temporal question of “when is this energy produced and consumed?”. Answering this requires modelling the hourly production against the hourly demand, and the present study bases its analysis on this hourly matching.

Because energy demand and generation in a city vary strongly from building to building and from hour to hour, Urban Building Energy Modelling (UBEM) has emerged as a standard workflow for capturing this variation (Reinhart & Cerezo Davila, 2016). This paradigm is based on bottom-up simulation: energy demand and production are modelled separately for each individual building and then aggregated at the neighbourhood or city scale. In this way the geometry, land use and hourly load profile of each building can be taken into account, giving a more accurate estimation of energy performance. The hourly generation-versus-demand analysis carried out in the present study falls within this bottom-up UBEM framework, enabling a building-by-building assessment of solar potential.

When PV systems are installed on a roof, not all of the energy produced is necessarily consumed on site, just as not all of the building’s electricity demand is covered by the panels. To evaluate and measure this two-sided interaction, researchers usually rely on two key indices: the Self-Consumption Index (SCI), which is the share of the produced energy that is used inside the building instead of being injected into the grid; and the Self-Sufficiency Index (SSI), which is the share of the building’s total demand that is supplied by its own solar system. These two indices depend strongly on the temporal match between production and consumption, and also on the capacity of any storage system, and they have been studied widely in the context of energy communities and the assessment of individual urban buildings (Todeschi et al., 2021). The present study also uses these two established indices to evaluate the real performance of the solar systems.

Most studies on self-consumption — including the city-scale assessment of Turin by Usta et al. (2025a, 2025b) — usually focus on the whole-city scale or on standard urban fabric: areas where roof space is large and where the production capacity of the panels is often matched to the energy demand, or even exceeds it. A dense historic centre, however, is a very different and far less studied case. In such a fabric, PV systems are usually undersized because of the lack of space, the roof geometry has oblique and unusual orientations, and the domestic demand peaks in the evening. These features mean that whole-city patterns do not transfer to these specific and complex areas. This study is carried out precisely to fill this gap: it provides an hourly, building-by-building analysis of self-consumption, developed specifically for the challenges of a dense historic centre — a case that previous large-scale studies had overlooked.

Chapter 3

From Aerial Photogrammetry to a Classified Point Cloud

3.1 Photogrammetry: from imagery to a point cloud

As already introduced in the state of the art, photogrammetry is a metric survey technique: it derives three-dimensional measurements from photographs through the central perspective projection, which establishes a one-to-one correspondence between a point on the ground and its image on the photograph (Kraus, 2007). Reconstructing an object therefore requires two sets of parameters for every image: the *interior orientation*, the internal geometry of the camera (focal length, principal point and lens distortions), and the *exterior orientation*, the position of the projection centre and the attitude of the camera at the moment of exposure. Once these are known, the collinearity condition — the object point, the projection centre and the image point lie on a single ray (Figure 3.1) — allows the corresponding (homologous) rays from two or more overlapping images to be intersected, fixing the 3D position of each point.

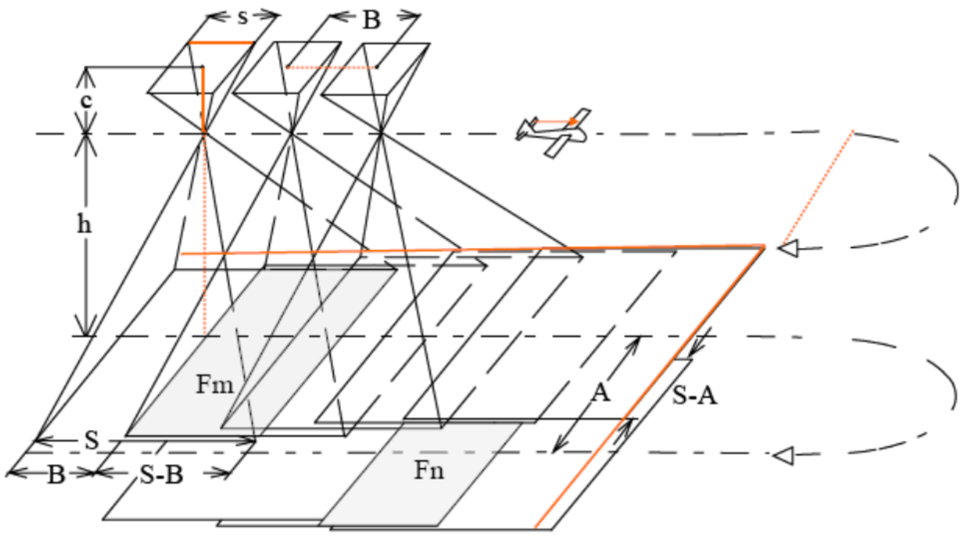

Figure 3.1: Aerial photogrammetric acquisition geometry and the ground sampling distance (GSD). Along the flight strip the consecutive exposures, spaced by the base B, image overlapping footprints on the ground; s is the image element and c the focal (principal) distance on the photograph, and S the corresponding ground dimension at flying height h. By similar triangles $s/c = S/h$, so the ground size of one pixel (the GSD) is $S = s\,h/c$. (Adapted from the Geomatics course material, Politecnico di Torino.)

In digital photogrammetry these steps are automated: the Structure-from-Motion (SfM) technique, implemented in photogrammetric software, extracts and matches large numbers of homologous points across the overlapping images (image matching; Chiabrando, Costamagna & Spanò, 2013) and solves the interior and exterior orientation of all the images at once in a single *bundle adjustment*, calibrating the camera in the same step (self-calibration) (Schönberger & Frahm, 2016). This first, sparse result — the tie-point cloud — is then densified by dense image matching, which computes a 3D position for almost every pixel by intersecting the homologous rays, giving the dense cloud used in the rest of this work.

Placing the model in a known reference system can follow two routes. In the classical *indirect* approach, *ground control points* (GCPs) of known terrain coordinates orient, scale and georeference the block, while independent *check points* (CPs) are kept aside to validate the result through their residual error. In *direct* photogrammetry, the coordinates of each projection centre are recorded in flight by a GNSS (Global Navigation Satellite System) receiver, so the exterior-orientation positions are known directly and the dependence on GCPs is reduced; the known positions both stabilise the bundle adjustment and provide the georeferencing (Kraus, 2007). The survey used here is of the direct type, as described below.

The imagery comes from the 2018 aerial photogrammetric flight of the Municipality of Turin, flown by aircraft (not by an unmanned aerial vehicle, UAV, as confirmed by the image metadata) with a Vexcel UltraCam camera. The full flight comprises 1324 photograms — each a TIFF larger than a gigabyte, with a near-infrared band also recorded — taken at a flying height of about 1670 m and a ground sampling distance of roughly 7.57 cm/pixel; every image is 14790×23010 pixels. The same flight also produced the municipal orthophoto and the digital surface model. From this block, the 135 nadir photographs covering the historic centre were used here. The photos overlap heavily, so the same part of the city appears in several of them, seen from slightly different positions.

These overlapping photos are processed using Structure-from-Motion (SfM). An image-matching algorithm searches for and identifies corresponding points across the overlapping images — for example the corner of a roof or a road marking — and links them from one photo to the next. Once these correspondences are found, a photogrammetric bundle adjustment recovers, at the same time, the position of each projection centre and the orientation of the camera for every photo. With the camera positions and orientations known, the corresponding rays from the different images can be intersected, and it is this intersection of rays that generates the point cloud. The name "Structure-from-Motion" reflects this: the 3D structure is recovered from the motion of the camera between shots (Schönberger & Frahm, 2016).

The result of this first step is a sparse tie-point cloud of 407012 points. It is called "sparse" because it only contains the strong, easily matched feature points (Figure 3.2).

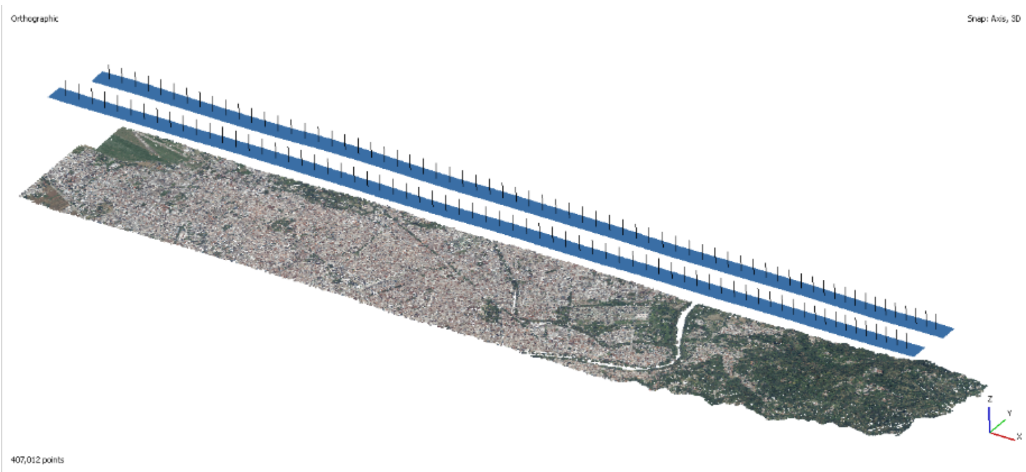

Figure 3.2: The sparse tie-point cloud (407012 points) with the camera positions, generated by Structure-from-Motion over the wider area.

A second step, dense matching, then goes back over the photos and computes a 3D position for almost every pixel, not only the strong feature points. This is what builds the much larger dense cloud used in the rest of the work. The area of interest was first selected with a region box (Figure 3.3).

Figure 3.3: Selection of the area of interest with a region box.

For the area of interest, the dense matching was run at ultra-high quality, producing a dense point cloud of 654664067 points (about 229 points per square metre on a flat surface), shown in Figure 3.4.

Figure 3.4: The full photogrammetric point cloud of the study area in the Turin historic centre (about 654 million points), coloured with the real photo colours.

3.2 Georeferencing accuracy

The aerial images were captured by a system that recorded the coordinates of each projection centre with an RTK (Real-Time Kinematic) GNSS receiver. Knowing the camera positions directly, rather than deriving them from ground control points (GCPs), is called *direct photogrammetry*. In the standard (indirect) workflow, GCPs measured on the ground are used before the dense cloud is built; here, because the projection centres are already known, these positions both optimise the photogrammetric adjustment and, at the same time, provide the georeferencing. The alignment still has to be validated, so its accuracy was checked against independent points. An orthophoto was also generated, with a ground sampling distance of 7.57 cm/pixel (Figure 3.5).

Figure 3.5: Orthophoto generated from the imagery (GSD 7.57 cm/pixel).

To estimate the error, independent check points were used. These come from the geodetic control network that the Municipality of Turin maintains for its topographic mapping and publishes on its geoportal: the *Punti Stabili di Riferimento* (P.S.R., stable reference points) and the *Punti Fiduciali* (P.F., cadastral fiducial points; Regione Piemonte, 2025), the same framework of known points from which the municipal 1:1000 map was produced (Figure 3.6). Their coordinates are of centimetric quality, so they are suitable as independent checks on the photogrammetric model. A note on the reference system is needed here: these points were originally given in the Italian national *Gauss–Boaga* datum, whereas the GNSS-based survey works in *WGS84* (World Geodetic System 1984); the check points were therefore transformed to WGS84 UTM 32N (Universal Transverse Mercator, zone 32 North; EPSG:32632, the cartographic projection used throughout this work) before the comparison (Figures 3.7 and 3.8). In total six check points were used.

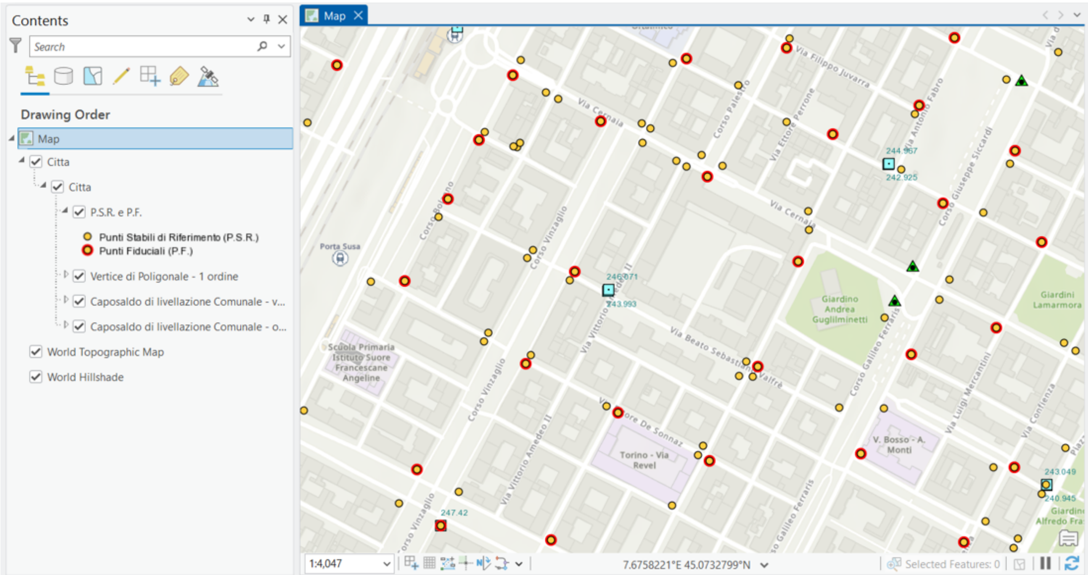

Figure 3.6: The Turin municipal geodetic control network on the geoportal: *Punti Stabili di Riferimento* (P.S.R.) and *Punti Fiduciali* (P.F.), together with the levelling benchmarks and traverse vertices. The check points used for the accuracy assessment were selected from this network.

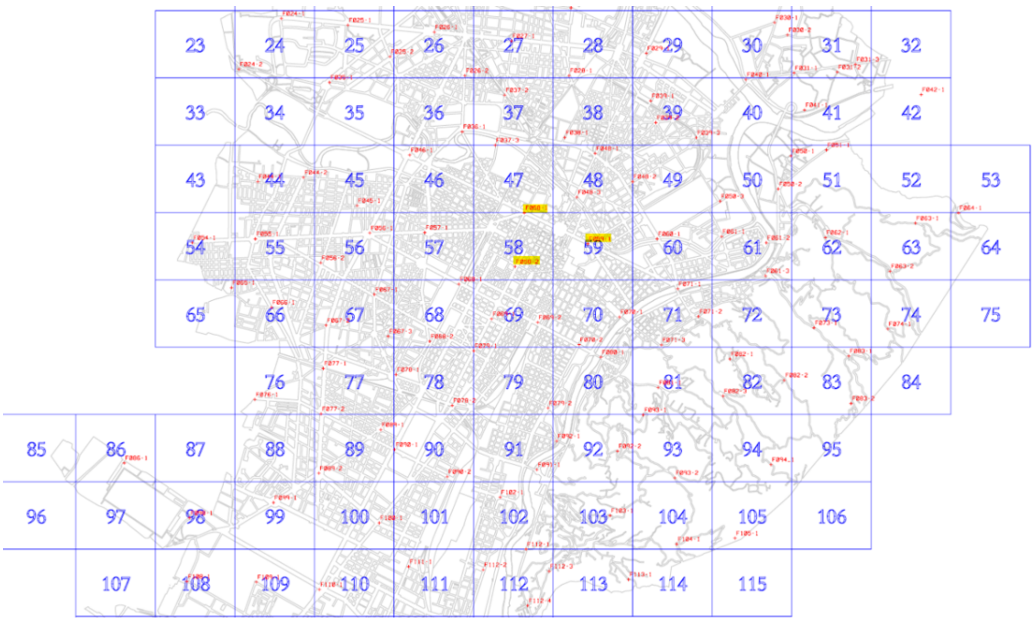

Figure 3.7: Check points used for the accuracy assessment (highlighted).

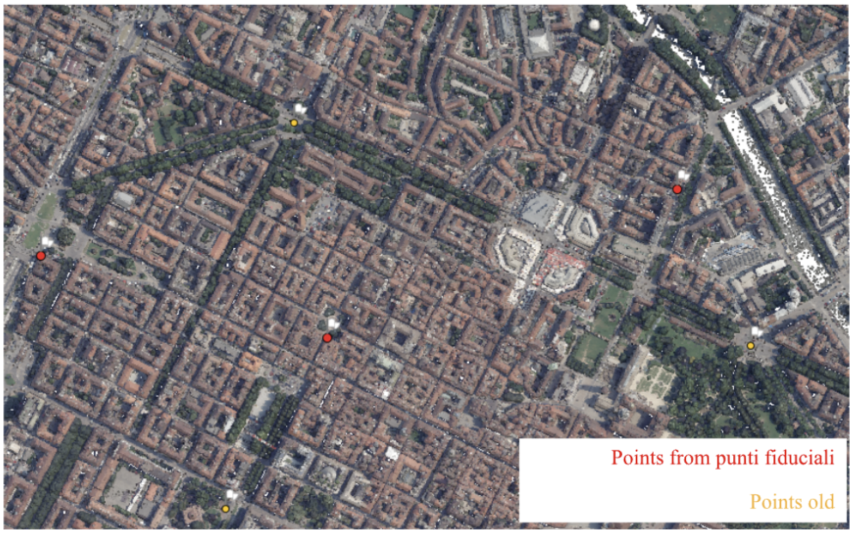

Figure 3.8: The three *punti fiduciali* selected from the Geoportale Piemonte dataset (Regione Piemonte, 2025), after coordinate transformation to WGS84 UTM 32N.

This accuracy assessment plays the same role as the classical check-point (CP) validation of the block orientation, and it was carried out on the oriented block before the dense cloud was generated. The resulting residuals for the cameras and the check points are reported in Table 3.1: the total root-mean-square (RMS) error over the check points was about 0.52 m, and the camera alignment error about 0.23 m. This camera-location error, from the RTK GNSS direct georeferencing, averages about 20 cm in X, 6 cm in Y and 10 cm in Z, so the planimetric accuracy (about 21 cm) is better than the vertical, as is expected for GNSS positioning. An accuracy of roughly 50 cm is adequate for the aims of this work; using more check points would make the assessment more robust.

Table 3.1: Georeferencing error report from the photogrammetric adjustment: the eight camera stations and the check points, with the total (RMS) error of each group. Coordinates are in WGS84 UTM 32N (EPSG:32632); the orthophoto ground sampling distance was 7.57 cm/pixel.

	Easting (m)	**Northing (m)**	**Altitude (m)**	**Error (m)**
Cameras				
11_1909	402720.34	4992083.90	1904.65	0.120
11_1910	402892.40	4992084.35	1904.76	0.074
11_1911	403064.36	4992084.99	1904.95	0.042
11_1912	403244.37	4992086.16	1905.10	0.035
11_1913	403416.28	4992088.18	1905.31	0.047
11_1914	403588.06	4992091.27	1905.27	0.086
11_1915	403768.04	4992094.60	1905.10	0.107
11_1916	403939.96	4992097.45	1904.88	0.153
Total camera error				**0.233**
Check points				
point 3	395225.11	4992284.4	247.62	0.477
point 2	395880.77	4992095.4	242.23	0.640
point 1	396682.56	4992437.4	226.21	0.277
F059_1	396849.67	4992078.7	230.00	0.693
F058_2	395648.34	4991702.8	242.86	0.577
F058_1	395803.75	4992588.9	237.92	0.268
Total check-point error				**0.516**

3.3 First classification attempt with Metashape

To build a clean 3D model, the high vegetation had to be removed, since the scene is a typical city landscape with trees, buildings, streets and flat ground. As a first attempt, the built-in classification tool of Metashape was used to separate the points into classes (Figure 3.9). Metashape classifies the ground automatically with a cell-based method: the cloud is divided into a grid, the lowest point in each cell is taken as a first ground estimate, a terrain surface is built from those points, and the remaining points are added to the ground class when their angle and distance to that surface stay below set thresholds. The other points are separated from geometric features, but in a dense photogrammetric cloud the features of trees and of building edges overlap, so this is where most of the classification errors appear.

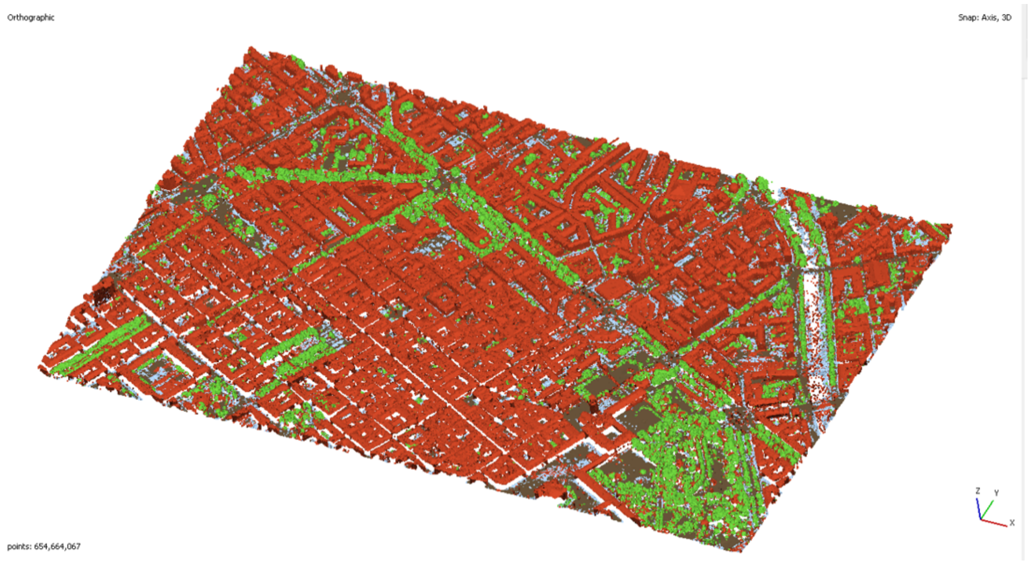

Figure 3.9: First classification attempt with the Metashape built-in tool (red = building, green = vegetation, blue = ground).

Using only the building and ground classes from this result to generate the DSM, without further cleaning, left a lot of vegetation wrongly labelled as building. As Figure 3.10 shows, this noise affects the surface model heavily.

Figure 3.10: DSM generated from the Metashape classification. Vegetation wrongly kept as building introduces strong noise into the surface.

Colouring the resulting surface by height makes the problem plain across the whole area: the roofs and streets read as smooth surfaces, but the tree-lined avenues and the green squares appear as a rough, speckled relief, because the vegetation kept in the building and ground classes is carried straight into the surface model as spurious bumps.

The Metashape classification was therefore not accurate enough: trees leaked into the building class and the resulting DSM was noisy. It was set aside, and a dedicated deep-learning model was trained to classify the points properly.

3.4 Classification with a deep-learning model

To classify the cloud properly, the TorchSparse++ library (Tang et al., 2022, 2023) was used to train a deep-learning model that labels every point into three classes: ground, building, and tree. Several areas were hand-labelled to train the model (Figure 3.11).

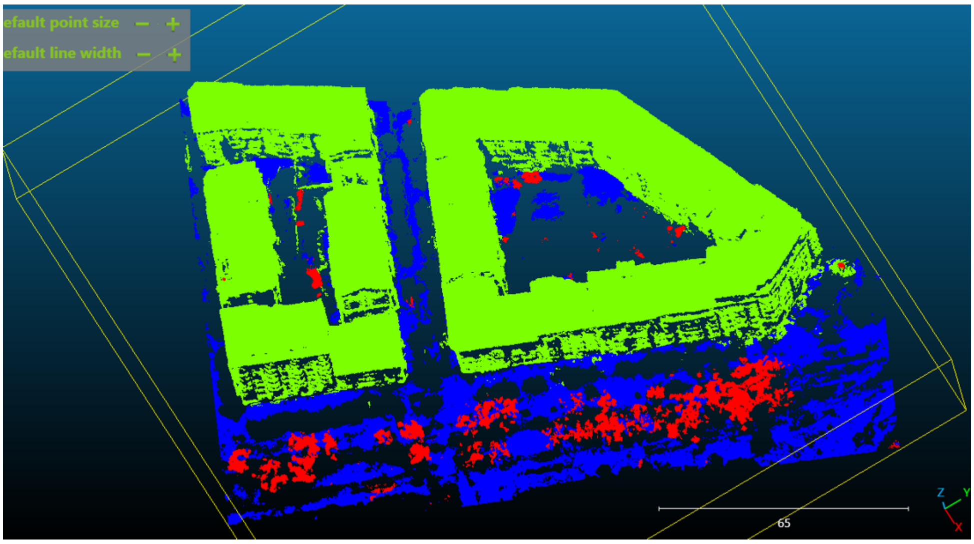

Figure 3.11: A hand-labelled training area. Each colour is one class: green = building, blue = ground, red = tree.

The labelled data consists of 8 tiles: 6 used for training and 2 kept for validation (to check how well the model works). Most tiles cover the dense building-block fabric of the historic centre. Because the tree class was otherwise very rare, one park area was split in half, one half placed in training and the other in validation, so that the model would see enough tree examples. A 1 m DTM (digital terrain model) was also built with the CSF (Cloth Simulation Filter) method (Zhang et al., 2016), which is used to give each point its height above ground (explained below). The tiles are summarised in Table 3.2.

Table 3.2: The 8 labelled tiles, their point counts and class balance, and their role (training or validation).

Tile (readable name)	Points	Class balance	Role
block_dense_A	3.77 M	building 94% / ground 6%	train
block_dense_B	5.99 M	building 95% / ground 4% / tree 1%	train
block_dense_C	4.18 M	building 87% / ground 12% / tree 1%	train
block_mixed_A	3.65 M	building 79% / ground 15% / tree 6%	train
block_mixed_B	5.57 M	building 77% / ground 15% / tree 8%	train
park_half_train	2.82 M	tree 65% / ground 35%	train
street_mixed_val	5.58 M	building 45% / tree 45% / ground 10%	validation
park_half_val	2.82 M	tree 58% / ground 42%	validation

The tiles were chosen in different parts of the study area so that the model would see a variety of urban situations: dense building blocks, open squares and streets, and the green area of the park. Their positions across the city are shown in Figure 3.12.

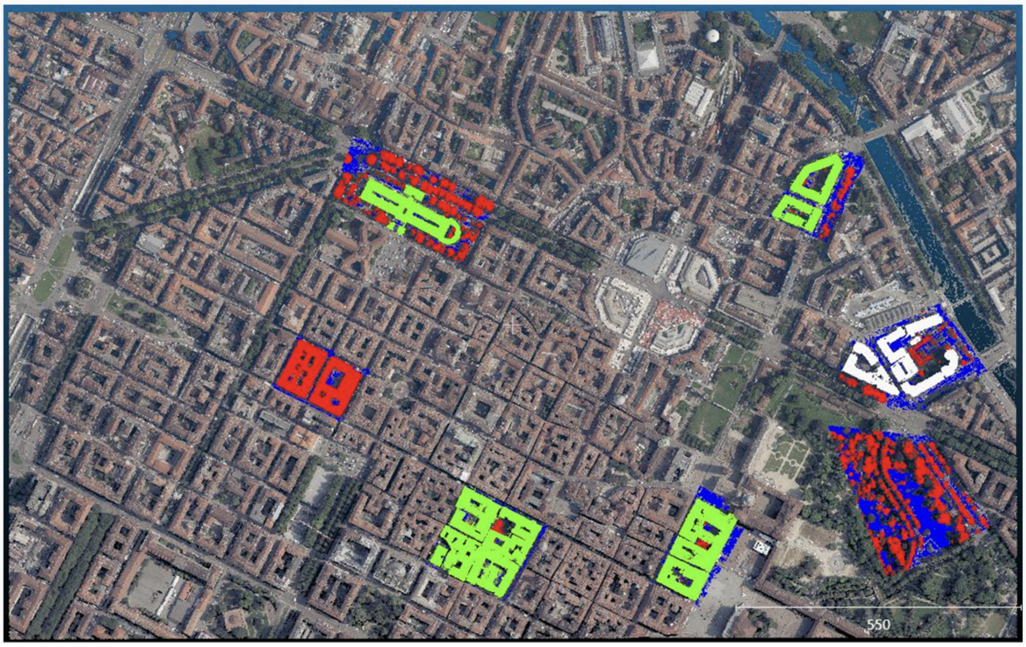

Figure 3.12: Location of the labelled tiles across the study area, shown on the aerial image. (Note: the colours here come from a different viewer than the other figures — in this image the dense building blocks appear in red, open ground in green, and vegetation in blue.)

3.4.1 Model input and the sparse representation

Each point is described to the model by 10 input features: its position within the crop (the local x, y, z), its true colour (the red, green and blue values carried over from the photographs), the three components of its surface normal, and its height above ground (HAG). The surface normal captures the local orientation of the surface around a point — roughly horizontal on a flat roof or the ground, tilted on a pitched roof, near-vertical on a wall — which is a strong clue for the class. The HAG is the single most useful feature: it states how high each point sits above the bare terrain, and so lets the model tell a roof from the street even when their colour is similar. It is obtained by subtracting the elevation of the 1 m CSF-derived DTM (stored as a GeoTIFF) from the point's own elevation.

The network takes a fixed-size input, so it is trained on small samples cut from the labelled tiles rather than on whole tiles at once: random 20 m × 20 m crops are taken from the tiles, and two things are done to each crop. First, the points are *voxelised* at 0.20 m: 3D space is divided into a regular grid of 0.20 m cubes (voxels), and all the points falling inside the same cube are merged into one input element. This fixes the spatial resolution the network sees and keeps each training sample at a manageable size, while 0.20 m is still fine enough to preserve roof edges and small structures. Second, because the crops are cut at random positions, the model sees the same urban fabric under many different framings, which acts as data augmentation and helps it generalise. Each voxelised crop is then passed through the network.

The important thing about TorchSparse is that it is a *sparse* library: empty voxels are simply skipped, and only the occupied ones are stored and processed. This is crucial for both memory and processing time, because a point cloud is mostly empty space.

3.4.2 Training and class imbalance

This kind of segmentation had earlier been attempted with the Minkowski Engine (Choy et al., 2019); TorchSparse++ worked noticeably better in practice, so the training was built around it. The model was trained for 80 epochs (handled by `train.py` and `dataset.py`), and the result improved epoch after epoch, so the best checkpoint came from the later stage of training. The final model reached about 98.5 % overall accuracy and 95.9 % mIoU on the validation tiles.

Because the study area is within the centre of Turin and the cloud comes from aerial imagery, most points naturally belong to buildings. This created an imbalanced-class problem: the tree class was very rare (less than 4 % at first). Class imbalance is a well-known difficulty in semantic segmentation, usually addressed either by re-weighting the loss towards the rare classes or by a focal loss (Lin et al., 2017). To handle it here, two strategies were used. First, *class weighting*, where the rare classes are given more weight in the loss, so the model is penalised harder for mislabelling them. Second, *minority rejection* (the `reject_minority` option in `dataset.py`), which forces a certain fraction of the training crops to contain enough tree points, so the model is pushed to actually see and learn the rare class instead of ignoring it.

3.4.3 Inference on the full cloud

Once the model was trained, it had to be run on the full cloud of 654 million points. This is too big to load into memory and classify in one go, so the inference was handled by a separate script (`inference_big_v2.py`). It first reads the whole file once and stores the features on disk, so the heavy file does not have to be read again. The area is then split into tiles, and inside each tile the same 20 m crops used in training are slid across with an overlap, so the model always has enough context and the result stays consistent at the borders.

Because the crops overlap, every point ends up being classified several times, from different crops. To get the final label, all these predictions are kept and a *majority vote* is taken for each point: the class predicted most often wins. This makes the result more robust than a single prediction, especially at the edges of the crops.

After running this over the whole area, all the unlabelled 654 million points are classified automatically (Figure 3.13).

Figure 3.13: The full area after automatic classification, shown in a different viewer: here buildings are orange, trees green, and the ground grey.

3.5 Evaluation

To measure how well the model works, it was evaluated on the validation tiles, comparing its predictions against the hand-labelled ground truth. The standard tool for this is the confusion matrix, which counts how many points of each true class were predicted as each class. From it the usual metrics are computed. Precision is the share of the points predicted as a class that really belong to it; Recall is the share of the true points of a class that the model recovered; the F1-score is the harmonic mean of the two; and IoU (Intersection over Union) is the overlap between the predicted and the true set of points divided by their union. Overall accuracy (OA) is the fraction of all points classified correctly, and the mean IoU (mIoU) is the IoU averaged over the three classes. The structure of the confusion matrix for a single class is shown in Table 3.3 (concept after Alfio et al., 2024).

Table 3.3: The structure of a confusion matrix for one class: true and false positives and negatives (concept after Alfio et al., 2024).

	Predicted: positive	**Predicted: negative**
Actual: positive	True Positive (TP)	False Negative (FN)
Actual: negative	False Positive (FP)	True Negative (TN)

The main evaluation was done on the balanced street tile (`street_mixed_val`), which is the only validation tile that contains all three classes in similar proportions, so its metrics are the most meaningful. The confusion matrix for this tile is shown in Table 3.4 (rows are the true class, columns are the predicted class).

Table 3.4: Confusion matrix on the balanced validation tile (counts). Rows = ground truth, columns = prediction.

true \ predicted	ground	building	tree	total
ground	543197	2061	1663	546921
building	40104	2446126	35070	2521300
tree	590	2789	2512873	2516252

From this confusion matrix the per-class metrics in Table 3.5 are obtained.

Table 3.5: Per-class metrics on the balanced validation tile. Overall Accuracy = 98.5 %, mean IoU = 95.9 %.

Class	Precision	Recall	F1-score	IoU
ground	0.930	0.993	0.961	0.924
building	0.998	0.970	0.984	0.968
tree	0.986	0.999	0.992	0.984
overall			OA = 0.985	mIoU = 0.959

The model is very strong on all three classes; the building and tree classes are almost perfect. The main confusion is that some building points (about 40000) are read as ground, which happens at the edge where the wall meets the pavement or where the roof eave meets the street.

The classified point cloud is the final product of this chapter and the input to the analysis that follows. From its building points a building-only digital surface model (DSM) was rasterised at a 5 cm grid, and this DSM is the surface on which the roof analysis of the next chapters is carried out. Two different quantities describe this surface and should not be confused: its *resolution* is 5 cm — the size of one raster cell, which fixes how finely the roof shape is sampled — whereas its *absolute accuracy* is the roughly 0.5 m georeferencing established above (Section 3.2), which fixes how precisely the whole model sits in real-world coordinates. A fine sampling grid and a coarser absolute position are not contradictory: the first is a matter of detail, the second a matter of placement.

Chapter 4

Building the 3D City (LOD2.1 Reconstruction)

4.1 Levels of detail

Once the building points were identified, the next step was to turn them into real 3D models. The level of geometric detail of a 3D building model is described by the LOD (Level of Detail) scale of the CityGML standard, which ranges from LOD0 to LOD4: LOD0 is only the 2D footprint, LOD1 a flat-top box, LOD2 adds the real roof shape, LOD3 adds façade details such as windows and doors, and LOD4 adds the interior (Figure 4.1). Biljecki et al. (2016) refined this coarse scale into finer sub-levels (for example LOD2.0 to LOD2.3), according to how accurately the building shape and the roof geometry are reconstructed. The standard itself has continued to evolve: CityGML 3.0 later revised the LOD concept and added an explicit space concept and a point-cloud representation (Kutzner et al., 2020).

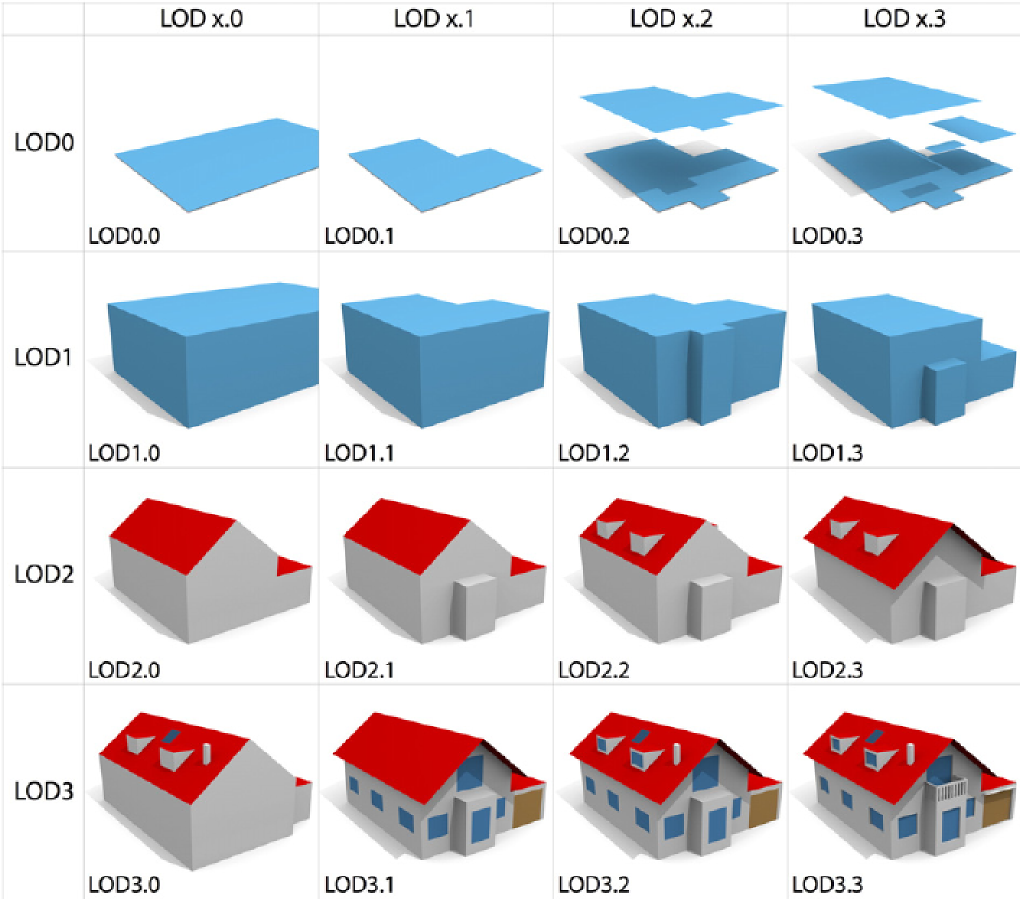

Figure 4.1: The refined Levels of Detail of Biljecki et al. (2016): each CityGML LOD (rows LOD0–LOD3) is divided into four sub-levels (columns x.0–x.3) of increasing geometric detail. In the LOD2 row the roof shape is reconstructed from LOD2.0 onward, whereas the roof superstructures (dormers, chimneys) are modelled only from LOD2.2. (Fig. 3 in Biljecki et al., 2016.)

The models produced here are LOD2.1 in this refined scale: they reconstruct the real roof slopes and shapes, with the walls extruded from the footprint, but the roof superstructures — dormers, chimneys and features of comparable size — are not modelled as separate geometry, which is precisely what would raise the model to LOD2.2. Those superstructures are instead handled later, face by face, by the roof-suitability classifier. As LOD2 models they also exclude façade details and roof openings such as windows, which belong to LOD3 and are outside the scope of this work.

4.2 Reconstruction with geoflow-roofer

The 3D models were created with geoflow-roofer, the open-source reconstruction tool developed by the 3D geoinformation research group at TU Delft, which underlies their national 3D BAG dataset of the Netherlands (Peters et al., 2022); the tool is freely available from the group's public repository. It was originally made for airborne LiDAR data, and here it was applied instead to the high-resolution photogrammetric point cloud from the municipal aerial flight, which is noisier and less regular than the clean airborne LiDAR the tool was designed for.

The inputs to the tool were the building points (taken from the classified point cloud) and the building footprints. The footprints come from the BDTRE (*Base Dati Territoriale di Riferimento degli Enti*), the shared geo-topographic database of the Piemonte region. The classified building points and the footprints used as input are shown in Figure 4.2.

Figure 4.2: Inputs to the reconstruction: the classified building points (slate) with the BDTRE footprint outlines (red), over the study area.

With the default settings, the reconstructed roofs came out spiky and inaccurate, because those defaults are tuned for the cleaner LiDAR case. The same tool was kept — only its configuration was changed, not its code — and the reconstruction was run many times with adjusted parameters until the models fitted the complex architecture of the centre of Turin as accurately as possible. In particular, the complexity factor was lowered (so the noise in the photogrammetric data would not create false roof planes), the plane-detection tolerance was loosened (so the noisier points would still be grouped into the correct roof plane), and the fallback limit was raised (so complex roofs would not collapse into a simple flat box). The full set of tuned parameters is given in Table 4.1. Changing these values is the intended way of using the tool and not a modification of the method: they are user-facing configuration parameters that geoflow-roofer exposes precisely so that the reconstruction can be adapted to point clouds of different density and noise, while the reconstruction algorithm itself is left unchanged. The values were therefore not chosen arbitrarily but tuned to the characteristics of the photogrammetric cloud, and their correctness is confirmed by the resulting geometry: the reconstructed LOD2.1 roofs reproduce the real roof planes and bring the solar-irradiation error down to about 3 %, against about 14 % for the flat-top model (Section 4.3).

Table 4.1: geoflow-roofer configuration used for the reconstruction, compared with the tool's LiDAR-oriented defaults.

Parameter	Value used	Default (LiDAR)
`complexity-factor`	0.45	0.888
`ceil-point-density`	10	20
`plane-detect-k`	35	15
`plane-detect-min-points`	100	15
`plane-detect-epsilon` (m)	0.55	0.30
`lod11-fallback-planes`	5000	900
`bld-class`	6	—
`grnd-class`	2	—
`lod13 / lod22`	true / true	—
`clip-terrain`	false	—

For each building the tool produces two levels: a simple flat-top version (LOD1.3) and a detailed version with the real roof shape (LOD2.1). Some buildings sat on the edge of the area and their point cloud was incomplete, so they were removed. In addition, the tool has a fallback: if a roof was too complex or too noisy to reconstruct as LOD2.1, it would fall back to the simpler LOD1.3, which happened for a small number of buildings. In the end, 1396 reconstructed buildings were obtained out of the original 1510 footprints, the large majority of them at LOD2.1.

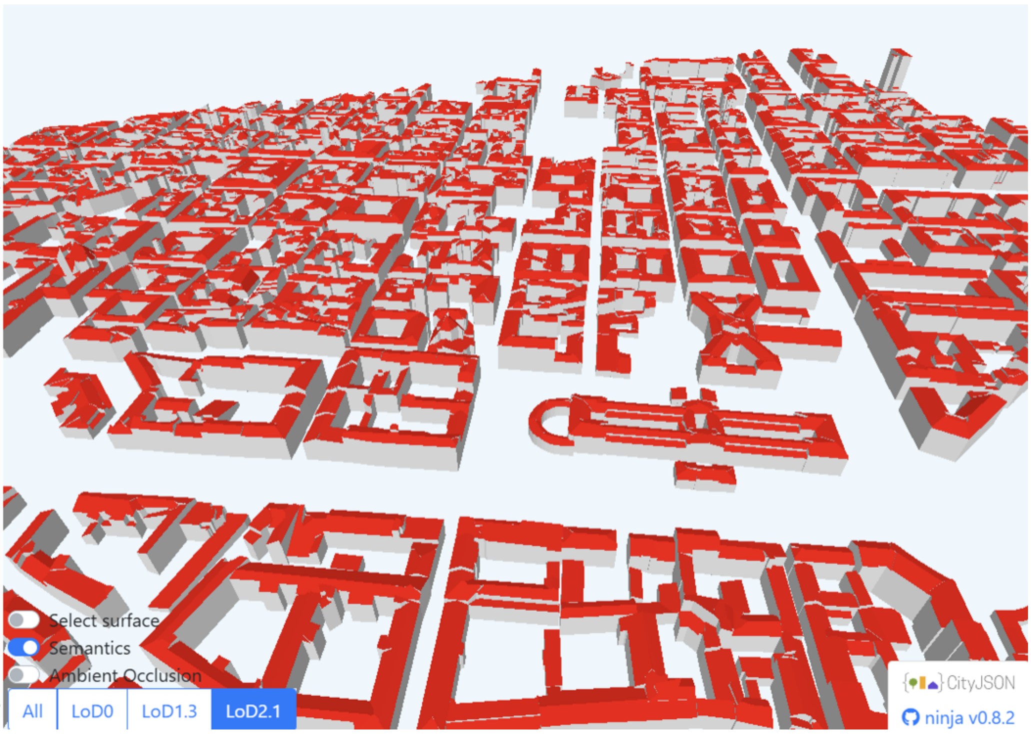

Figure 4.3: The reconstructed study area at LOD2.1 (real roof shapes), shown in ninja, the free web viewer for CityJSON.[1] Roof surfaces are red, walls grey.

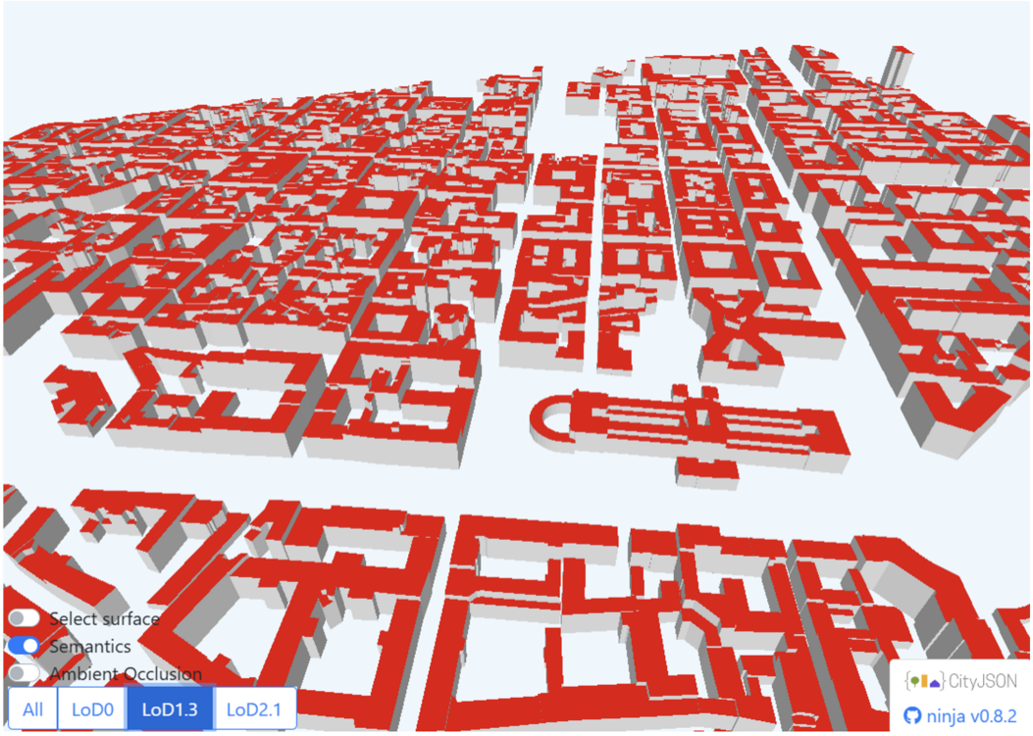

Figure 4.4: The same area at LOD1.3 (flat-top boxes), for comparison. The roof shape information is lost at this level.

4.3 Why the roof shape matters

The difference between the two levels is clearest on a single building. Figures 4.5 and 4.6 show the Anagrafe building, one of the more distinctive buildings of the centre, reconstructed at LOD2.1 and at LOD1.3.

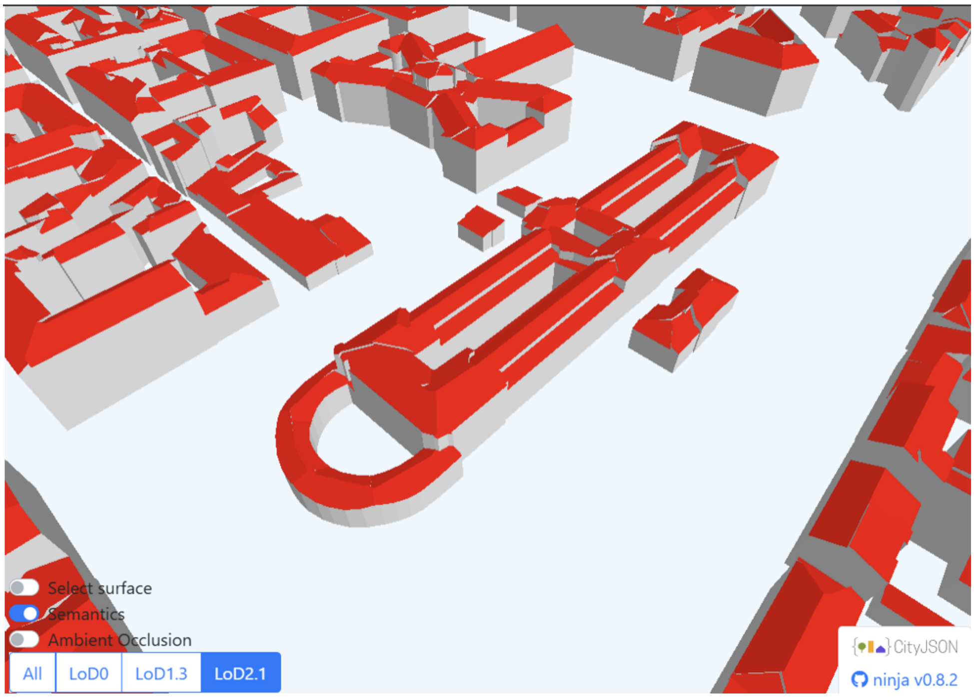

Figure 4.5: The Anagrafe building at LOD2.1: the real roof slopes and faces are reconstructed.

Figure 4.6: The same building at LOD1.3: a flat top, with no slope or roof-face information.

The roof shape matters for the solar calculation: the flat-top LOD1.3 model gives about 14 % error in the estimated irradiation, while the detailed LOD2.1 model brings it down to about 3 %. The reason is physical: a solar-irradiation calculation needs the slope and the orientation (aspect) of each roof surface to know how much sun it receives over the year. The LOD2.1 model reproduces the real roof faces, each with its own slope and orientation, so a south-facing pitch and a north-facing pitch are treated differently, as they should be. The flat-top LOD1.3 model, by contrast, replaces the whole roof with a single horizontal surface that has neither slope nor orientation, so it cannot tell a well-exposed pitch from a poorly-exposed one, and the irradiation estimate is wrong by a correspondingly large margin.

4.4 Limitations of the reconstruction

The large majority of buildings were reconstructed well, but the most architecturally complex buildings — those with internal courtyards, curved or circular structures, rooftop gardens and glazed roofs — were sometimes reconstructed with fragmented or broken roof surfaces (Figure 4.7). Two factors combine to cause this. First, the geometry itself is genuinely difficult: these roofs depart strongly from the simple planar shapes that the reconstruction assumes, so the plane-fitting struggles to represent them. Second, there is a mismatch between the two input sources: the photogrammetric point cloud is precise and follows the real roof shapes, while the BDTRE footprints used to delimit each building are coarser and simplified (Figure 4.8). The BDTRE is a regional dataset, but its building footprints are derived from a simplified version of the 1:1000 municipal map, with attributes that refer to the 1:2000 scale, so their level of detail, although the most detailed available, is not high enough for a building-by-building analysis such as this one. Where the footprint does not match the real roof outline, the points are clipped to the wrong boundary and the plane-fitting degrades. These cases are a minority, but they show that the method depends not only on the quality of the point cloud but also on the agreement between the cloud and the footprint data.

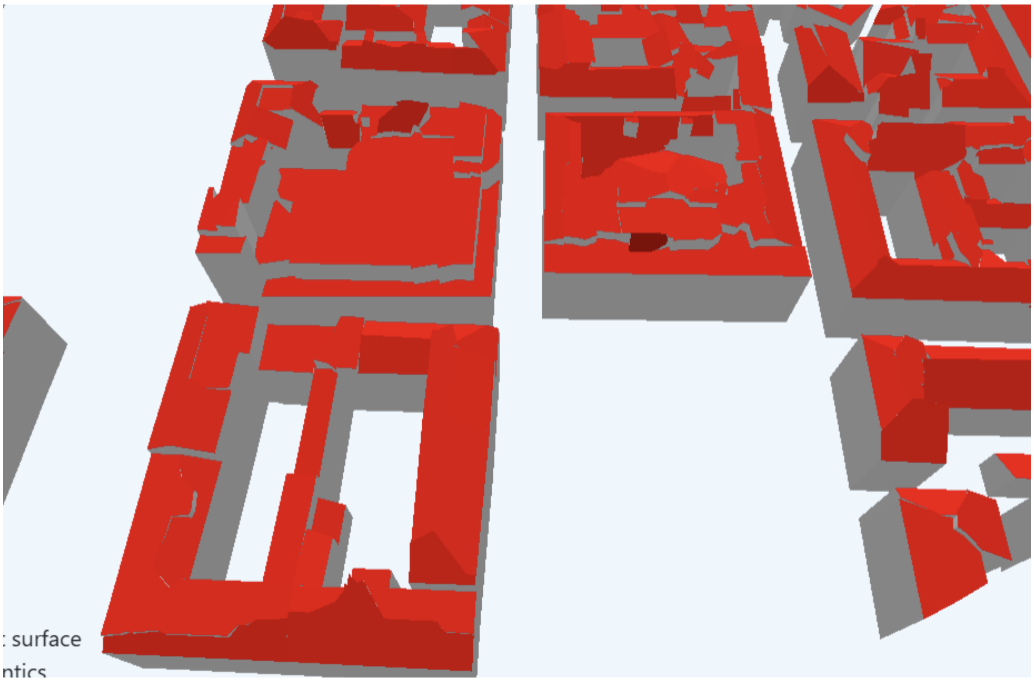

Figure 4.7: Example of degraded reconstruction on architecturally complex buildings: the roof surfaces are fragmented and broken.

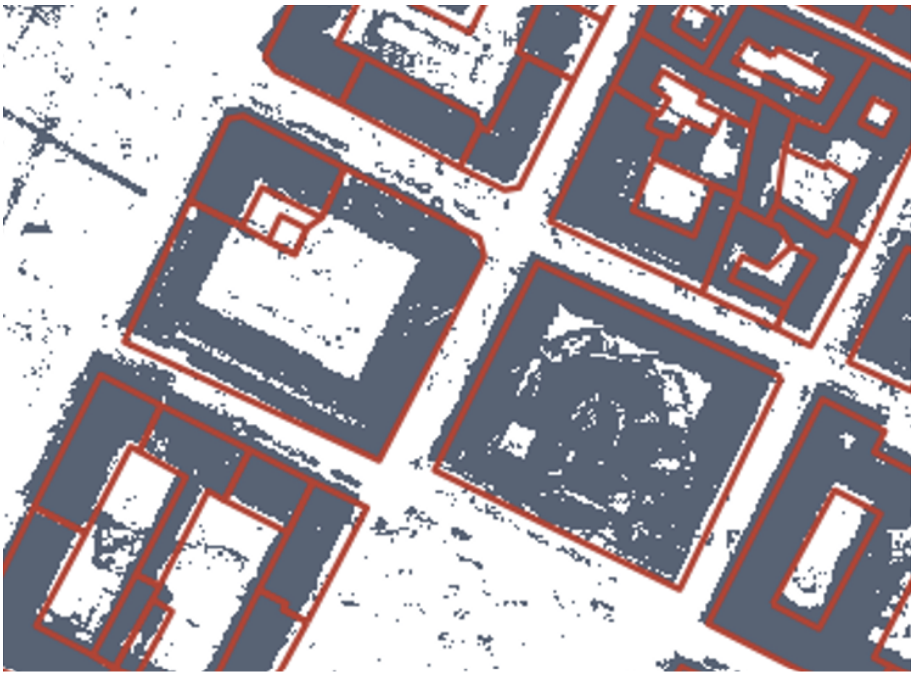

Figure 4.8: A cause of the problem: the precise building points (slate) do not always match the coarser BDTRE footprints (red), so the points are clipped to the wrong boundary.

Chapter 5

Finding the Usable Roof Area

The segmentation step produced a vector layer of roof segments for the whole study area, each segment coloured by the direction it faces (Figure 5.2). Starting from these segments, this chapter turns the roofs into a usable-area result: the solar irradiation on the roofs is computed first; then a classifier decides, segment by segment, whether a surface is a good roof; that decision is merged onto the LOD2.1 roof faces to give the usable area of every roof; and finally an interactive viewer is presented for exploring the reconstruction and the classification. The next chapter then uses this usable area to test the 0.35 coefficient and to model the hourly self-consumption.

By *usable roof area* is meant the part of each roof that could realistically carry photovoltaic panels. A roof surface is counted as usable only when it is a genuine, continuous roof plane, oriented and inclined well enough to receive sun, and large enough to hold at least one panel; surfaces that are too small, or that are taken up by obstructions — windows, skylights, chimneys and the raised *abbaini* (dormers) described later in this chapter — are excluded. These criteria are applied in two stages: a trained classifier first decides, segment by segment, whether a surface is a good roof, and a short set of physical rules then enforces the size limits, in particular a minimum area of about $3\,\mathrm{m}^2$ (the footprint of a single panel), below which a fragment is too small to be used, and an upper size above which a detected obstruction is too large to be a real window or chimney. This matters because the usable area is the quantity that the following chapter converts into installable PV capacity and into the self-consumption of the study area.

The segmentation itself followed the roof-segmentation method of Usta et al. (2025), applied here to the Turin data and at a finer resolution. Starting from the 5 cm building DSM — whose absolute accuracy follows the roughly 0.5 m georeferencing established in the previous chapter — an aspect raster (the compass direction each pixel faces) and a slope raster were derived using the raster terrain-analysis tools. The aspect raster was then segmented with the `i.segment` tool of GRASS GIS, which merges neighbouring pixels of similar aspect into one region. The difference threshold of 0.2 set in that method (their Figure 2) was used: on the tool's scale from 0 to 1 this is a low value, so only pixels facing almost the same direction are merged into one segment. This keeps separate roof pitches and small features apart, which matters at the 5 cm resolution. The regions were then turned into polygons with the Polygonize tool, and each polygon was given its median slope and aspect. Following Usta et al. (2025a), a polygon was classed as flat when its slope was below 8°, and otherwise placed in one of twelve 30° orientation bins that divide the compass: besides the four cardinal directions (north, east, south and west), each 90° quadrant holds two further intermediate directions, giving twelve orientation classes in all. The steps are summarised in Figure 5.1.

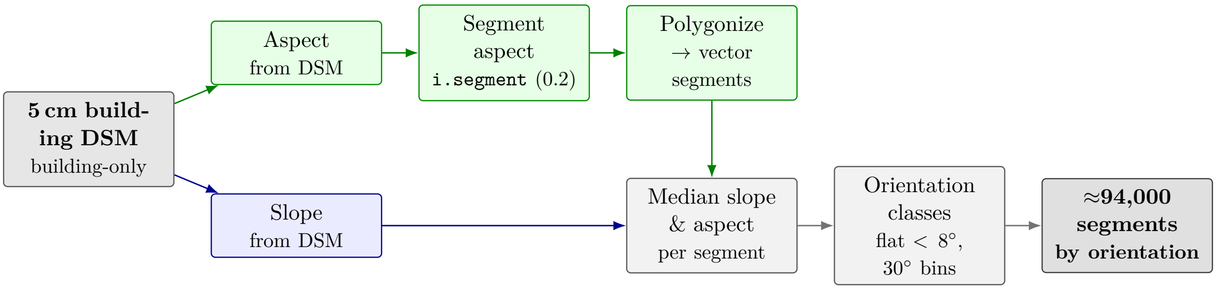

Figure 5.1: The roof-segmentation workflow. Green boxes use the aspect raster, the blue box uses the slope raster, and the grey boxes are the joined result: from the 5 cm building DSM to ≈ 94,000 vector roof segments grouped by orientation.

Figure 5.2: The vectorised roof segments over the study area, coloured by orientation class: a flat class for near-horizontal roofs, plus twelve 30° sectors around the compass (here 0° points south, -90° east and 90° west). The twelve sectors cover the four cardinal directions and two intermediate directions within each quadrant, so each colour is one of the twelve orientation classes. This is the segmentation result that the following steps work on.

5.1 Solar irradiation on the roofs

To know how much sunlight reaches each roof, the annual solar irradiation was computed with the r.sun model (Šúri & Hofierka, 2004), the open-source solar-radiation tool of GRASS GIS, run on a 1 m building DSM. This follows the GIS-based rooftop-solar workflow developed for Turin by the Politecnico research group (Usta et al., 2025a; Todeschi et al., 2021). The inputs were the diffuse and global radiation from PVGIS (2023), the Linke turbidity from Meteonorm (2023), and a ground albedo of 0.25, the value normally taken for dense urban areas in that workflow. This gave an annual irradiation value (SI_Ann, in kWh/m^2 per year) for every roof segment. These values are used later, in the area comparison, to turn the suitable area into a gross energy figure (Layer 3).

5.2 Classifying the roof segments

After the segmentation step there were around 94575 polygons, each one a connected group of pixels sharing the same aspect bin. Most of these were not real usable roof faces. In Usta et al. (2025a) all segments smaller than $6\,\mathrm{m}^2$ were simply deleted at this point, but their DSM had a resolution of 1 m while the one used here was 5 cm. At 5 cm resolution, real roof features such as dormers — in Italy called *abbaini*, the small raised structures that carry a vertical window for the inhabited space under a pitched roof — and other small faces produce many small polygons that would be wrongly thrown away by a fixed area cut. All segments were therefore kept, and instead a classifier was trained to decide, for each polygon, whether it was usable or not.

In total 259 polygons were hand-labelled directly in QGIS, on the vectorised DSM. Each was given one of three labels: good roof (75 polygons), bad surface (108 polygons), and window or obstruction (76 polygons). The third class was important because windows and skylights appear in the segmentation as small raised patches, and the model had to learn to detect them: a roof covered in skylights should not be counted as usable area.

For each polygon the classifier uses three groups of numbers. The first group describes the shape of the polygon: its area, its perimeter, how many vertices it has per square metre, and how compact and solid it is. Noise patches and windows tend to be small and jagged, while real roof faces are larger and smoother. The second group carries the slope and aspect values from the segmentation step, with the aspect encoded so that north-facing values near 350° and 10° are treated as close rather than opposite. The third group compares each polygon to the other polygons on the same building: how different its slope is from the building median, how much its facing direction deviates from the building average, and how small it is relative to the other segments on that roof. This last group was the strongest signal for picking out windows and chimneys, because they stand out from the roof faces around them even when their own shape looks reasonable.

An XGBoost model (Chen & Guestrin, 2016) — a widely used, open-source gradient-boosted decision-tree library — was trained on the 259 labelled polygons. This roof-suitability classifier is an original contribution of this thesis: whereas the segmentation follows Usta et al. (2025a), the classifier and its feature set were developed here and trained on data labelled specifically for this work. The model was evaluated with five-fold cross-validation, so that every labelled polygon was scored by a model that had not seen it during training. The weighted F1 score was 0.923.The per-class results are given in Table 5.1 and the confusion matrix in Table 5.2.

Table 5.1: Five-fold cross-validated classification report on the 259 labelled polygons.

Class	Precision	Recall	F1	Support
good (1)	0.897	0.933	0.915	75
bad (2)	0.923	0.889	0.906	108
obstruction (3)	0.948	0.961	0.954	76
overall (OA)			0.923	259
macro avg	0.923	0.928	0.925	259
weighted avg	0.923	0.923	0.923	259

Table 5.2: Confusion matrix from the five-fold cross-validation (rows = true class, columns = predicted class).

true \ predicted	good (1)	bad (2)	obstruction (3)
good (1)	70	5	0
bad (2)	8	96	4
obstruction (3)	0	3	73

Most of the confusion is between good (1) and bad (2). This is expected, because both are roof surfaces and the difference between a usable roof and a poor one is not always sharp. The result that matters most for this use case is that good (1) and obstruction (3) are never confused with each other: a roof is never read as a window, and a window is never read as a roof. This is important because only the good class is later turned into usable area, so a roof-window mix-up would feed straight into the final result.

The trained model was then applied to the full dataset of 94575 polygons across the whole study area. The result is shown in Figure 5.3: green polygons are classified as good roof (class 1), red as bad surface (class 2), and blue as windows or obstructions (class 3). Only the good polygons carry forward to the area calculation in the next step. A close-up of a few blocks (Figure 5.4) shows this in detail: the cyan obstruction patches — windows, skylights and chimneys — are scattered across the red and green roof faces, numerous but individually small.

Figure 5.3: The classifier applied to the whole study area. Green = good roof (class 1), red = bad surface (class 2), cyan = window or obstruction (class 3). Obstructions are many in number but small in area, so at this scale they appear as scattered dots rather than large patches.

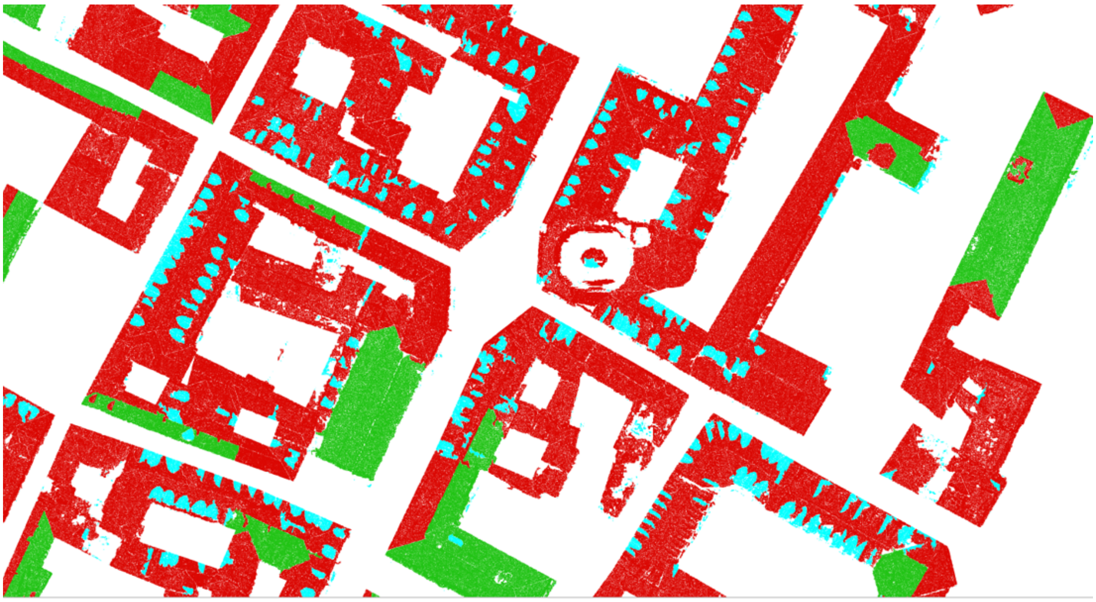

Figure 5.4: A zoom on a few blocks of the classifier result. At this scale the individual obstructions are clear: the cyan patches are the windows, skylights and chimneys detected as class 3, sitting on the red (bad) and green (good) roof faces. This is the level at which the roof-suitability decision is made, face by face.

After the model predicted a class for every polygon, two simple physical rules were applied on top of the raw predictions. First, anything predicted as an obstruction but larger than $8\,\mathrm{m}^2$ is too big to be a real window or chimney, so it is reassigned to good or bad. Second, any good polygon smaller than $3\,\mathrm{m}^2$ is too small to hold a panel, so it is moved to bad.

A known error: large abbaini (dormers)

The classifier is not perfect. The clearest error type is shown in Figure 5.5: three large *abbaini* — the raised dormer structures, common on the pitched roofs of the historic centre, that carry a vertical window to light an inhabited space directly under the roof (they are roofed structures, not glazed openings, and the window sits on their vertical face) — were labelled as good roof when they should not be. This happened because the 76 hand-labelled obstructions were all small windows and chimneys, so the model never saw raised structures of this size. Their large footprint also meant that the $8\,\mathrm{m}^2$ obstruction rule reassigned them away from the obstruction class. Cases like this are a minority, but they show that the training set should ideally include large abbaini and other large raised structures as well.

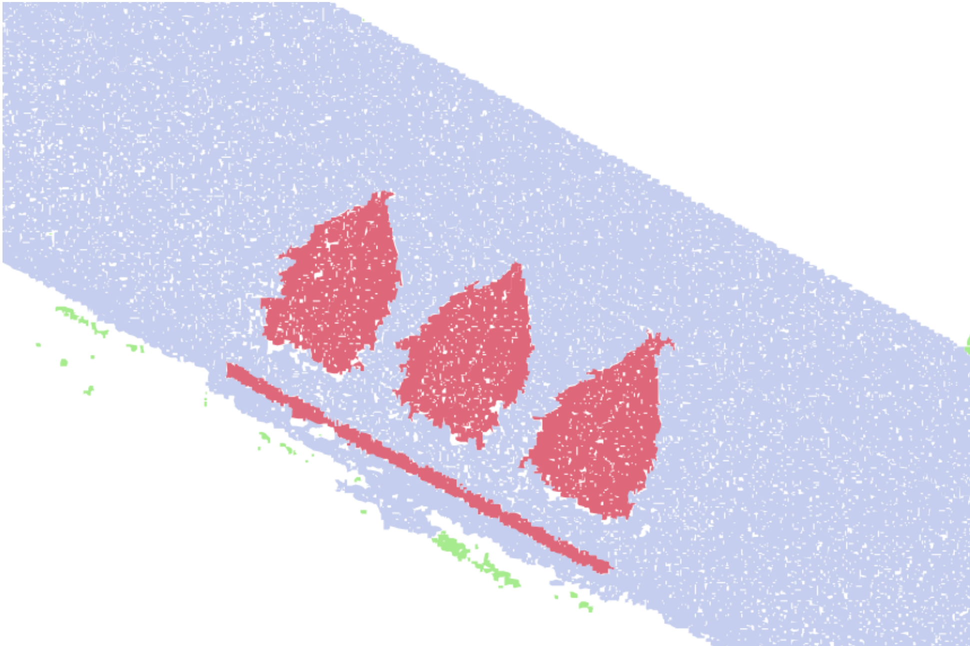

Figure 5.5: A known error: three large *abbaini* (dormers, red) — raised structures that carry a vertical window for the inhabited space under the pitched roof — were classified as good roof, because the training set only contained small windows and chimneys.

5.3 Merging the classification onto the LOD2.1 faces

Once the classifier had assigned a label to each segment, those labels had to be connected to the LOD2.1 roof faces. The two datasets do not line up: the classifier worked on segments cut from the 5 cm raster, while the LOD2.1 model has 5434 planar faces that were built independently by geoflow-roofer from the same point cloud. A single segment can cross several faces, and a single face can have many segments sitting on top of it.

Before the merge, the segments were cleaned of noise: polygons smaller than $0.5\,\mathrm{m}^2$ or with an extreme number of vertices per square metre (pixel-rasterisation speckles) were dropped. This reduced the 94575 polygons to 67898 that were carried into the merge.

To connect the two datasets a spatial intersection was used. For each LOD2.1 face, every classifier segment that overlapped it was found and the exact area of each overlap was computed. In this way a segment that crosses three faces contributes to all three, in proportion to how much of it falls inside each one. At the end, every face knows how many square metres of its area were classified as good, as bad, and as obstruction. The LOD2.1 face layer is shown in Figure 5.6, and the result of laying the classification over those faces is shown in Figure 5.8.

Figure 5.6: The LOD2.1 roof faces over the study area: 5434 planar faces reconstructed by geoflow-roofer. Each face is the unit onto which the classification is merged.

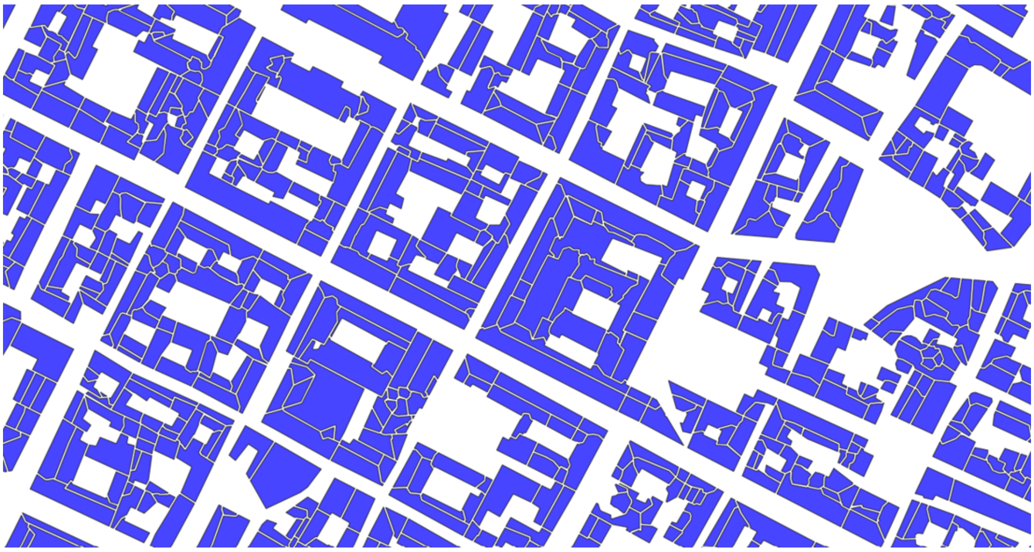

Figure 5.7: Detail of the LOD2.1 faces, showing how each roof is divided into separate planar faces.

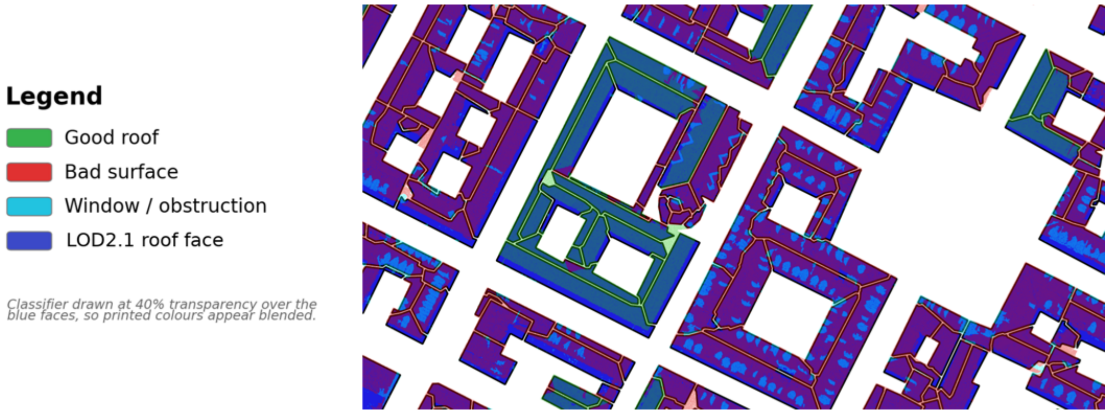

Figure 5.8: The roof-suitability classification (good = green, bad = red, obstruction = cyan) overlaid on the LOD2.1 roof faces (blue) at 40 % transparency; the legend gives the source colours. The un-blended classification is shown in Figure 5.3.

The result of this step is a new face layer in which each of the 5434 LOD2.1 faces carries its good area in m^2, its bad area, and its obstruction area. Only the good area goes forward into the comparison.

5.4 An interactive viewer for the results

To make the results easier to explore, a small interactive 3D web viewer that runs in a browser was built. It shows the reconstructed study area and can colour the roofs by any of the results. Figure 5.9 shows it coloured by the per-face classifier verdict: dark green faces are suitable for PV, light green are partially suitable, purple are obstructed, red are unsuitable, and grey have no coverage. The colour can be switched to other results too – the suitable area, the gross energy, the annual irradiation, or the ratio between the 0.35 coefficient and the suitable area, which is the main comparison of this work.

Clicking a building opens a side panel (Figure 5.10) with all of its values in one place: the cascade bars (0.35 coefficient, LOD1.3 effective area, Layer 1 and Layer 2), how much the coefficient over- or under-estimates that building, the gross Layer 3 energy, and context values such as the built area, the bounding box, the roof type, the number of roof faces, the LOD2.1 deviation percentage and the reconstruction error. Clicking a single roof face shows that face's area, slope, orientation and whether it is flat, its classifier verdict with the good, obstructed and coverage percentages, and the face's own annual irradiation and Layer 1/2/3 contribution. In the example shown, the selected building has more suitable area than the coefficient would predict – the coefficient under-estimates it by 40 % – the opposite of the city-wide average, and another reminder that the coefficient cannot be trusted building by building.

The viewer is an original contribution of this thesis, built specifically for this work with the open-source deck.gl and MapLibre GL libraries, which draw the roof faces as interactive WebGL layers over a base map in the browser. It does not compute anything new; it only displays the results produced by the previous steps. In the LOD2.1 view the walls are not drawn: a rendering problem meant they did not show correctly, so the roof faces appear as floating surfaces. This is only a limitation of the viewer display and does not affect the analysis, which works on the roof faces. The simpler LOD1.3 view shows the flat-top blocks with their walls, and for the full LOD2.1 mesh with walls the viewer links out to the ninja CityJSON viewer (Vitalis et al., 2020).

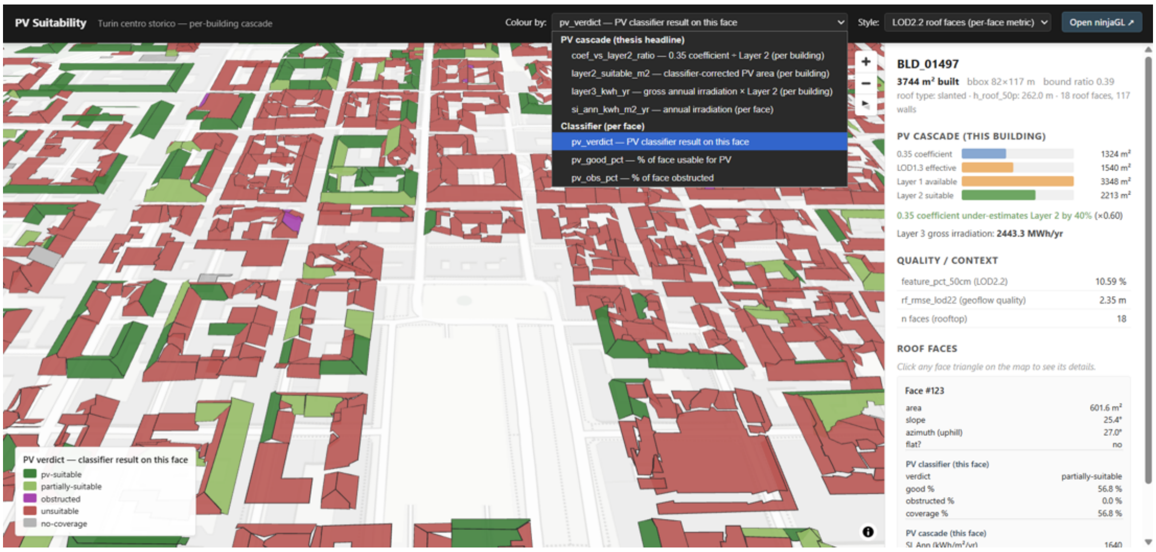

Figure 5.9: The interactive web viewer, with roof faces coloured by the classifier verdict (green = suitable, light green = partially suitable, purple = obstructed, red = unsuitable, grey = no coverage). The side panel shows the full PV cascade for the selected building and the details of one roof face.

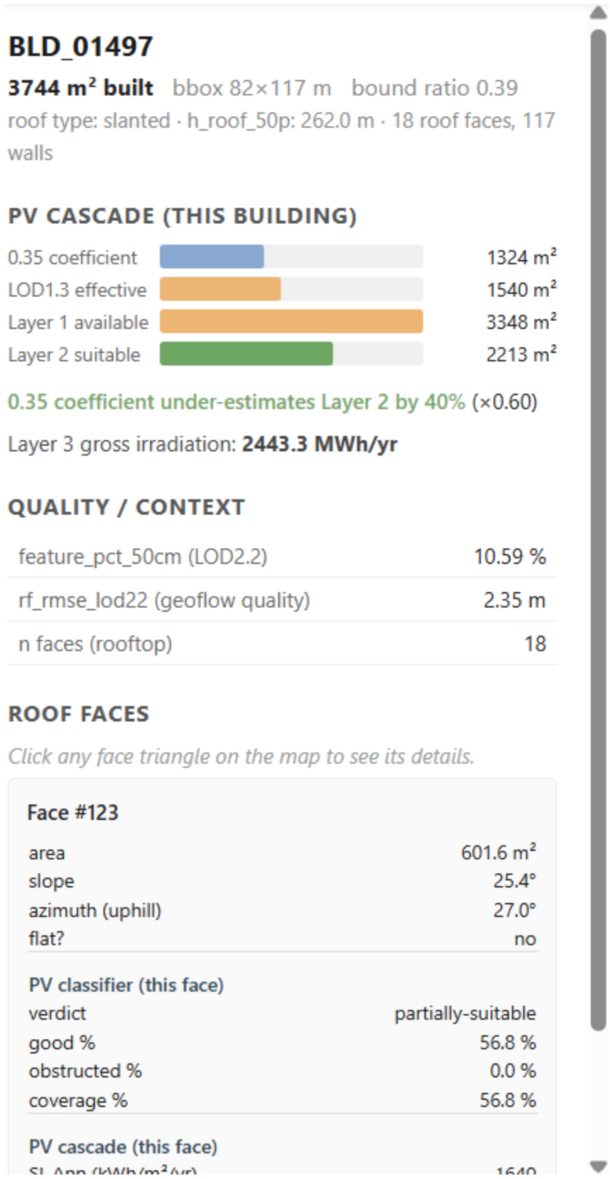

Figure 5.10: The side panel for a selected building, showing the four-layer cascade, the coefficient comparison, the gross Layer 3 energy, the quality and context values, and the details of a single roof face.

Chapter 6

Rooftop PV Potential and Self-Consumption

The previous chapters reconstructed the roofs of the study area and measured, face by face, the area that is genuinely usable for photovoltaics. This chapter turns that geometry into use. It first compares the measured usable area with the simpler building models normally used for city-scale PV estimates — above all the 0.35 coefficient — to see how reliable that shortcut is for the single building. It then models the timing of the energy: for the residential buildings, how much of the electricity a rooftop system would produce is actually consumed on site, hour by hour. From this point on, the analysis is about the energy use of the buildings. The building-by-building test of the 0.35 coefficient and the hourly, per-building self-consumption analysis for a dense historic centre are the original contributions of this part; the underlying tools are existing and openly available — the r.sun irradiation model, the ARERA (Italy's energy, networks and environment regulator) residential demand profiles, the ISTAT (Italian national statistics institute) census, and the urban building energy modelling method of Usta et al. (2025b).

6.1 Comparing the building models for PV area

One of the questions raised in supervision was how the standard fixed coefficient compares with a detailed, per-building estimate of the rooftop area available for PV. Two approaches were therefore compared on the same set of buildings. The first is the 0.35 coefficient, which simply multiplies the building footprint by 0.35. The second is the detailed LOD2.1 model, which carries the real roof slopes and a geometry close to reality, and was enriched with the result of the classifier — the suitable area.

On the LOD2.1 model the estimate was built up in layers, each one closer to the real usable area (Table 6.1 and Figure 6.1). Starting from the gross LOD2.1 roof area, the parts that are not clean roof (chimneys, dormers and similar) were removed to obtain the available area (Layer 1). From the available area, only the faces that the classifier marked as good roof were kept, which gives the suitable area (Layer 2). Finally, multiplying the suitable area of each face by its annual solar irradiation gives a gross energy figure (Layer 3).

To find the obstructions for Layer 1, the per-face deviation was used between the measured roof surface (the DSM) and the reconstructed LOD2.1 face. Where the DSM departs from that face – chimneys, small dormers, and noise left over from the reconstruction – the surface is not clean roof. The share of each face taken up by these parts is the *feature percentage*, and Layer 1 is the face area with that share removed.

Table 6.1: City-wide rooftop PV estimates from the different methods, summed over the 1396 reconstructed buildings.

Estimate	City-wide total
0.35 coefficient (footprint × 0.35)	292952 m^2
LOD2.1 gross (all faces)	833751 m^2
Layer 1 – available (LOD2.1 minus obstructions)	576347 m^2
Layer 2 – suitable (classifier)	202796 m^2
Layer 3 – gross energy upper bound	210.9 GWh/yr

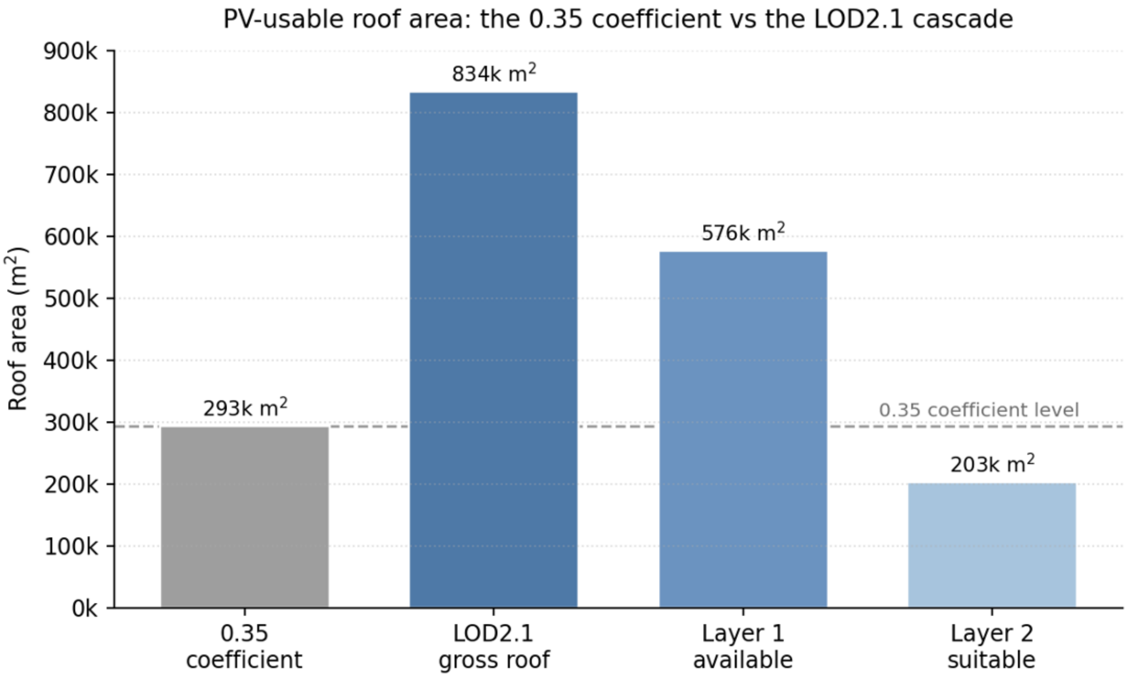

Figure 6.1: The PV-area cascade across the 1396 reconstructed buildings, from the 0.35 coefficient down to the classifier-suitable area (Layer 2).

Out of the 1396 reconstructed buildings, only 714 had at least one roof surface that the classifier marked as suitable for PV installation, giving a classifier-suitable area (Layer 2) of only 202796 m^2. Compared with this, the 0.35 coefficient over-estimates the usable roof area by 44.5 % across the whole centre, and this aggregate figure hides much larger per-building errors. Looking at single buildings (Figure 6.2), the coefficient is unreliable in both directions: the ratio between the coefficient area and the suitable area ranges from 0.70 to 2.42, so for some buildings it is far too high and for others too low. The error is worst on flat roofs (Figure 6.3): for the few flat-roofed buildings in the study area the coefficient over-estimates the suitable area by more than twenty times, while for the pitched roofs that make up almost all of the study area it is close to correct at the median.

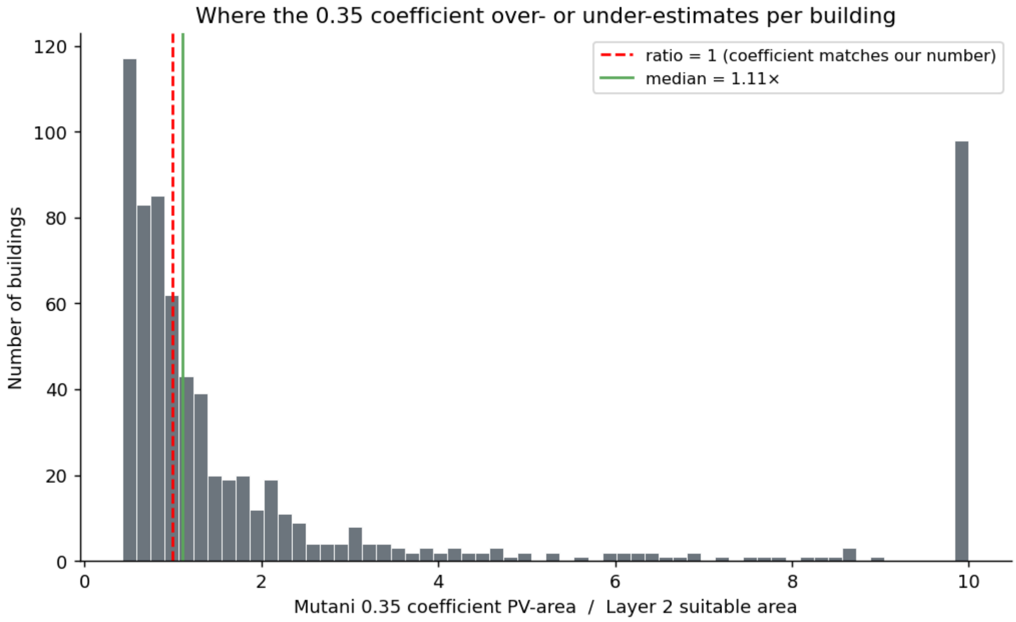

Figure 6.2: Per-building ratio between the 0.35 coefficient area and the classifier-suitable area, over the 714 buildings with suitable roof. A ratio of 1 means the coefficient matches the classifier number; the spread shows it is too high for many buildings and too low for others.

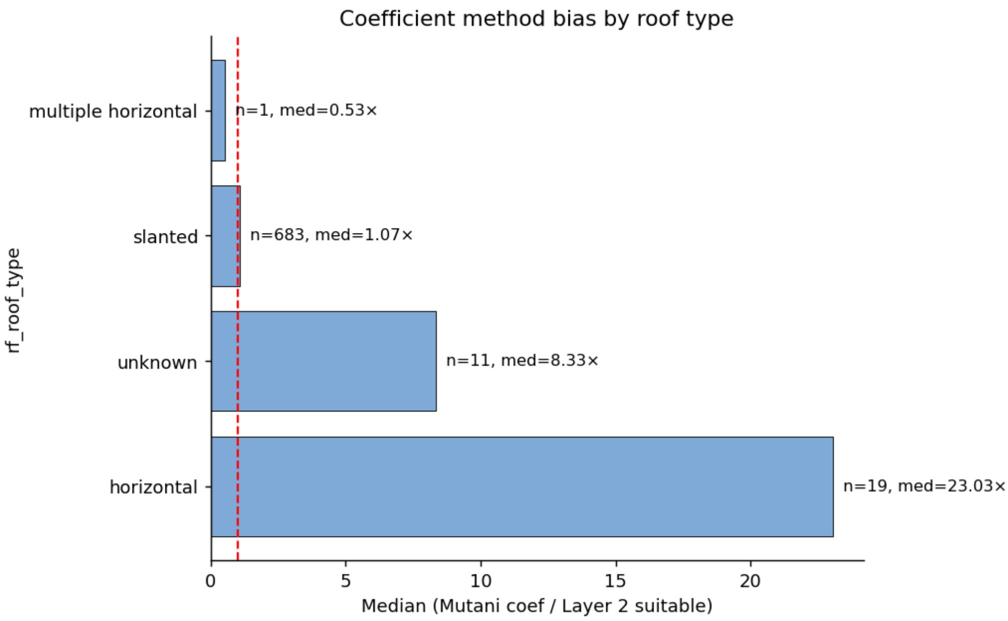

Figure 6.3: Median ratio (coefficient area / classifier-suitable area) by roof type. The coefficient is close to correct for slanted roofs (683 buildings) but over-estimates flat roofs heavily; note the small number of flat-roofed buildings (19).

Layer 3 should be read as a gross upper bound. It is the annual irradiation (computed earlier with r.sun) falling on the suitable area, with no system losses or performance factors applied, so the energy a real system would generate is lower.

6.2 From producible energy to hourly self-consumption

The previous chapter found how much roof area in the study area is suitable for PV and how much energy it could produce. That is a yearly total, but it does not say whether the electricity is produced at the same time as it is needed. A roof can produce a lot of energy at noon while the people in the building use most of their electricity in the evening. This part of the chapter looks at the timing: for the residential buildings, the hourly PV production is compared with the hourly electricity demand over a whole year, and how much of the production is used on site is measured.

Two indices describe this, following the same definitions as Usta et al. (2025). The Self-Consumption Index (SCI) is the share of the PV production that is used inside the building instead of being exported to the grid (the public electricity network). The Self-Sufficiency Index (SSI) is the share of the building's demand that is covered by its own PV. A high SCI means the panels are well matched to the demand; a high SSI means the building depends less on the grid. For each hour the self-consumed energy is the smaller of production and demand; what is left over is either exported (production above demand) or imported (demand above production).

6.3 Which buildings, and how many families

This analysis only concerns residential buildings, because it compares PV production against household electricity demand. Starting from the 1396 reconstructed buildings, those whose use is residential (including mixed-use buildings with apartments) were kept, and garages and anomalies removed (footprint below $40\,\mathrm{m}^2$ or height below $3\,\mathrm{m}$). This left 882 residential buildings. Of these, the self-consumption analysis covers the 417 buildings that have both a PV-suitable roof area (from the classifier) and a resident load.

The number of families in each building is needed to size the demand. The ISTAT 2021 census was used, which gives the number of resident families per census block; that number was shared among the residential buildings of the block in proportion to their volume, following the method of Usta et al. (2025b). The 417 buildings hold 7915 families in total.

6.4 PV production

The hourly PV production of each roof face was computed with the formula

$$P = \mathrm{PR} \cdot H \cdot (f \cdot A) \cdot \eta,$$

where the performance ratio is $\mathrm{PR} = 0.80$, the module efficiency is $\eta = 0.24$ (premium monocrystalline), A is the classifier-suitable area of the face (Layer 2, the classifier-suitable area, Section 6.1), and H is the solar irradiance on the face for that hour. The factor $f = 0.80$ is the active-area fraction: it accounts for the inactive frame around each module, the border that carries no cells and produces no electricity. Only this frame factor is applied, and not the more common 0.6 coefficient, because the 0.6 also bundles in a reduction for obstructions and unusable roof area — and that reduction has already been carried out, face by face, inside the Layer 2 classifier area. Applying the full 0.6 on top of Layer 2 would remove the unusable area a second time. One small loss the method does not capture separately is the packing loss on a clean face – the spacing left between panel rows and for maintenance – which neither Layer 2 nor the frame factor removes. The production of all the faces of a building is then summed to give the building's hourly production.

The hourly irradiance H combines two sources. Its hourly *shape* over the day and year comes from PVGIS (a typical meteorological year for Turin), and its monthly *magnitude* is rescaled to match the irradiation measured on each face from the DSM in the previous chapter. This keeps the real shading of each roof by its neighbours, instead of assuming an open sky. H is therefore a hybrid: its shape comes from PVGIS and its magnitude is the value measured in this work.

6.5 Residential demand

The electricity demand of each building was built from the number of families and a typical residential demand profile. The hourly shape comes from the ARERA "Prelievo medio orario" measured data for resident domestic customers (ARERA, 2023), with separate shapes for weekdays, Saturdays and Sundays. The annual demand of each family was set to 2700 kWh per year, the ARERA *cliente tipo* reference, which is also the value used by Usta et al. (2025) so that the results can be compared.

Two limits of this demand model should be stated plainly. First, the 2700 kWh per family is a reference value, not a measurement: the ARERA measured average per customer is lower (about 1030 kWh per year), and the choice of 2700 raises the demand and therefore affects both indices; it is used here for comparability with the reference study. Second, the month-to-month variation of the demand uses a placeholder seasonal pattern (a winter and a summer peak) rather than measured ARERA monthly data, because the monthly series was not available at the time; this mainly affects the seasonal balance, not the annual totals.

6.6 Production versus demand over the day

Figure 6.4 compares the average hourly PV production with the residential demand on a typical winter day and a typical summer day. In both seasons the demand has its highest peak in the evening, around 20:00, when the PV production has already stopped. This evening peak is therefore always covered by the grid. In winter the midday production rises to roughly the level of the midday demand, while in summer it goes well above it and the surplus is exported. The self-consumed energy is the overlap between the two, shaded in the figure.

The green band in this figure is easy to misread, so it is worth one clarification. The self-consumed energy was not read off the two average curves. It was worked out building by building and hour by hour: for each building and each hour the smaller of its own production and its own demand was taken, and only then averaged over all the buildings to draw the green band. This matters because the buildings are not all alike. At midday some roofs make more than the home needs and send the surplus to the grid, while others make less; the part sent to the grid can never count as self-consumed. For this reason the green band sits a little below the level where the average production and the average demand appear to overlap. Taking the overlap of the two average curves instead would hide these differences between buildings and report more self-consumption than really happens, so the per-building calculation is the correct and more honest one. It follows the same definition of self-consumption used by Usta et al. (2025b).

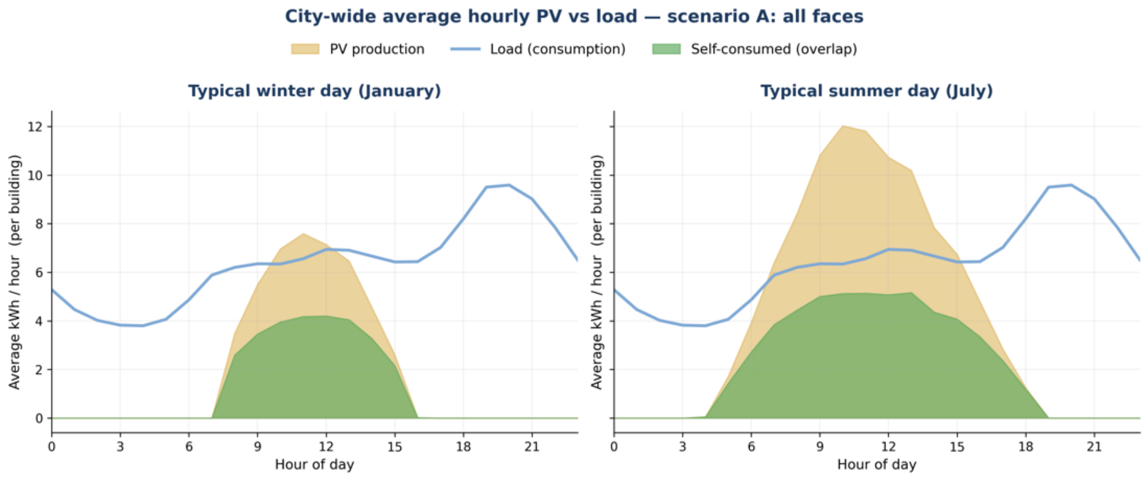

Figure 6.4: Average hourly PV production and residential demand per building on a typical winter day (left) and summer day (right), all-faces scenario. The green area is the self-consumed energy (the overlap); the demand peaks in the evening, after production has ceased.

6.7 The effect of roof orientation

The historic centre has an unusual roof geometry that matters for PV. Because the orthogonal Roman street grid of the centre is rotated away from the cardinal axes and runs along the intercardinal directions, the roof pitches face mostly the intercardinal directions (SE, SW, NE, NW) rather than the cardinal ones (N, S, E, W): about 73 % of the faces point to an intercardinal direction, and true south-facing roofs are rare. Figure 6.5 shows the measured monthly irradiation per face by orientation, which follows the expected order (south and south-east highest, north lowest) and so confirms the shading-aware values.

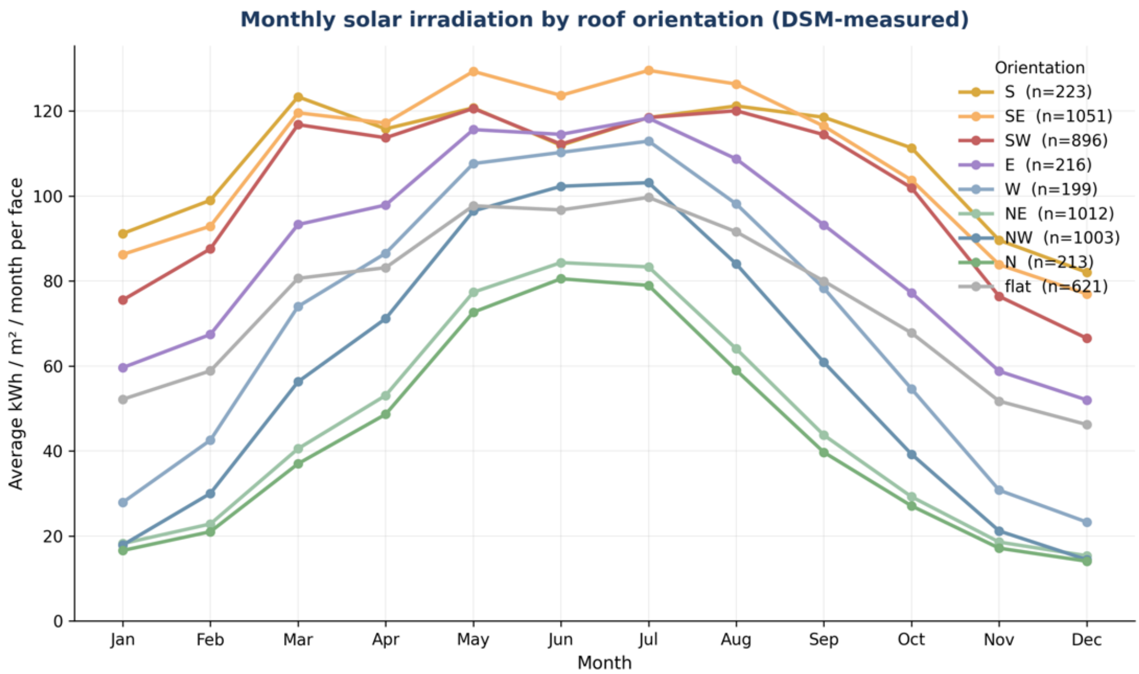

Figure 6.5: Monthly solar irradiation per roof face by orientation (DSM-measured). The number of faces in each orientation class is given in the legend.

This geometry decides where the city's PV production comes from (Figure 6.6). The south-east faces alone provide about 32.9 % of the total annual production, and together with the south-west faces they dominate, while the true-south faces contribute little simply because there are so few of them. The all-faces scenario (A) keeps every orientation, including the north-facing faces, which receive the least irradiation and are the least convenient for PV; the south-only scenario (B) restricts the estimate to the well-exposed south-facing and flat faces, in line with the usual practice of installing rooftop PV only on favourably oriented roofs (Mutani & Todeschi, 2021).

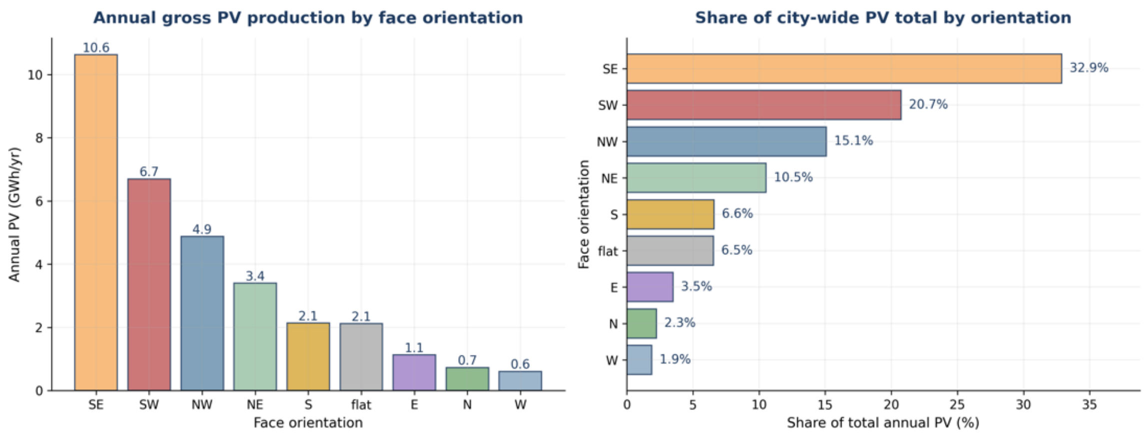

Figure 6.6: Annual gross PV production by face orientation (left) and each orientation's share of the city total (right). South-east dominates because of the diagonal street grid, not because those faces are individually the best.

6.8 Self-consumption results

Two scenarios are reported. In scenario A every suitable roof face is equipped with PV. In scenario B only the south-facing faces (S, SE, SW and flat) are equipped. The city-wide results are given in Table 6.2.

Table 6.2: City-wide self-consumption results for the residential buildings, for the all-faces scenario (A) and the south-only scenario (B), with the PV reference scenario of Usta et al. (2025b).

Quantity	A: all faces	B: south only	Usta et al. (2025b) PV
Residential buildings	417	335	—
Families	7915	6671	—
Annual PV (GWh/yr)	11.05	7.43	—
Annual load (GWh/yr)	21.37	18.01	—
PV / load sizing ratio	0.52×	0.41×	> 1 (oversized)
City-wide SCI	51.76 %	55.49 %	63.12 %
City-wide SSI	26.76 %	22.88 %	55.47 %
Median per-building SCI	70.52 %	76.77 %	—
Median per-building SSI	31.66 %	30.32 %	—

The per-building distributions are shown in Figures 6.7 and 6.8. The two ways of reporting each index mean different things. The city-wide SCI of 51.76 % is an aggregate (the total self-consumed energy divided by the total production over all buildings), and it is dominated by the few large producers. The per-building *median* SCI is higher, 70.52 %, because a typical small building consumes a larger share of its own modest production. The same holds for the SSI: an aggregate of 26.76 % against a per-building median of 31.66 %. A clear ceiling appears in the SSI distribution: no residential building reaches an SSI above about 48 %, so on its own rooftop PV never covers more than about half of a building's electricity.

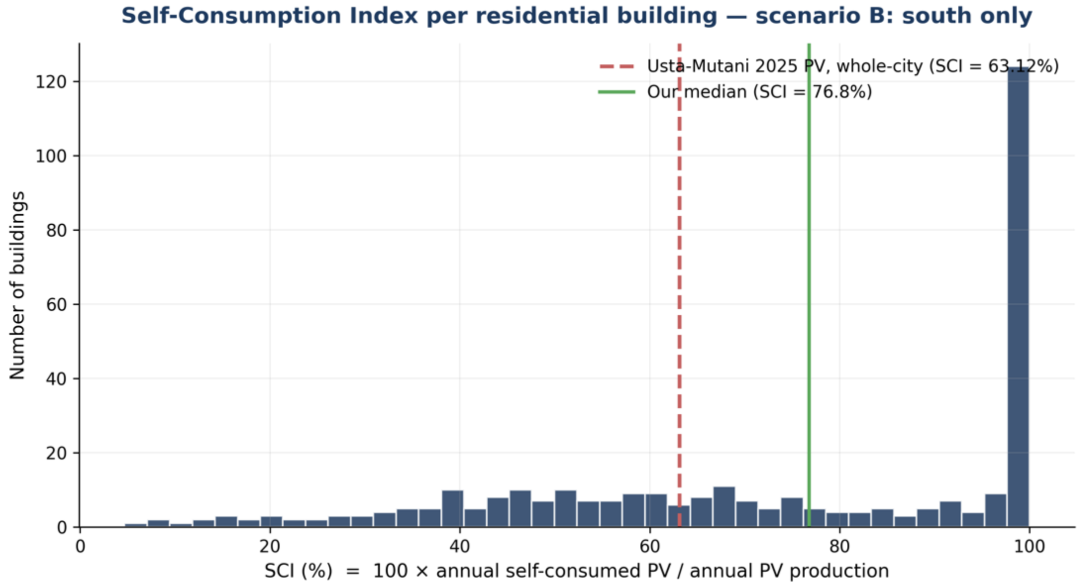

Figure 6.7: Self-Consumption Index per residential building, all-faces scenario. The green line is the per-building median (70.52 %); the dashed line is the whole-city PV scenario of Usta et al. (2025b).

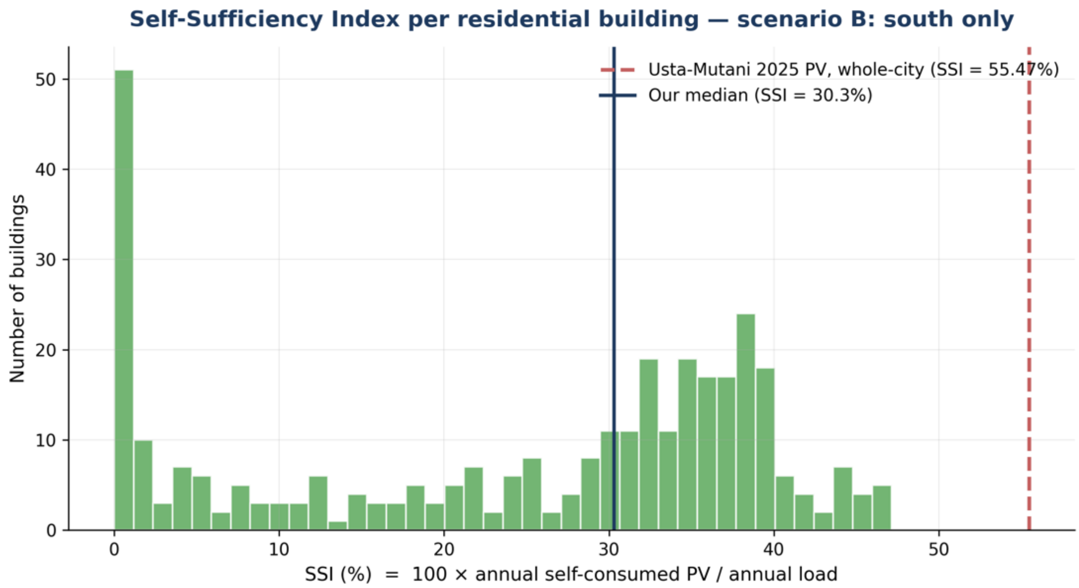

Figure 6.8: Self-Sufficiency Index per residential building, all-faces scenario. No building exceeds an SSI of about 48 %.

Restricting the panels to the south-facing faces (scenario B) raises the SCI slightly (city-wide from 51.76 % to 55.49 %) but lowers the total production, the load served and the SSI, because fewer faces are used. The comparison of the two scenarios across all the headline quantities is shown in Figure 6.9.

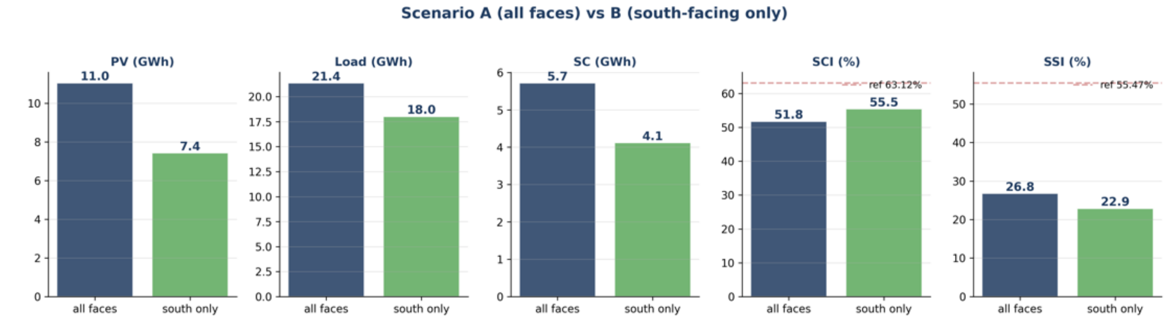

Figure 6.9: All-faces (A) versus south-only (B) scenario across the headline quantities. The dashed lines mark the Usta et al. (2025b) PV reference.

6.9 Discussion: comparison with the city-scale Turin study

Usta et al. (2025b) report a PV-electricity scenario for the whole city of Turin with a self-consumption index of about 63 % and a self-sufficiency index of about 55 %. The study-area results here are below these values (all-faces SCI 51.76 %, SSI 26.76 %). This is not a contradiction: the two studies describe different areas. Their figures are for the whole city, where roofs are larger and many face south or are flat, and where the PV is well matched to the demand. The historic centre is the opposite case. Beyond describing a different, denser part of the city, the study area is also subject to historical and cultural constraints — notably the heritage protection of the built environment (the Italian Cultural Heritage Code, D.Lgs. 42/2004) — which can limit or forbid rooftop PV on protected buildings and set it further apart from a whole-city assessment; these constraints are not applied to the suitable area here and are a natural direction for future work. Two clarifications are needed for a fair comparison: their value is a whole-city aggregate, so it should be read against the aggregate here (not the per-building median), and their demand is the whole-city residential electricity (about 1000 GWh) against the 21 GWh, so their number is a reference rather than a strict benchmark.

The reason the centre sits lower is the sizing ratio. The active classifier-suitable roof area produces only about 52 % of the annual residential demand (a PV/load ratio of 0.52×): the study area is PV-undersized. With production this scarce relative to annual demand, one might expect almost all of it to be self-consumed, yet the city-wide self-consumption index is only 51.76 % (the per-building median is higher, 70.52 %). The shortfall comes from the daily mismatch between supply and demand: solar production peaks around midday, when residential demand is low, so the midday surplus is exported to the grid and, without storage, cannot be recovered later. This timing mismatch is what caps the self-consumption. The self-sufficiency is then limited by the same two factors – the scarce annual production and the evening demand peak, which falls after production has stopped – so rooftop PV covers only about a quarter to a third of the residential electricity. In short, a substantial share of the rooftop PV is used on-site, but it can still supply only a limited part of the demand; the binding constraints are the available roof area and the daily timing of supply and demand, not the self-consumption efficiency.

Chapter 7

Conclusions

This thesis set out to measure, building by building, the usable rooftop photovoltaic (PV) area of a dense historic centre, starting from aerial photogrammetry, and to test whether the fixed coefficient normally used for city-scale estimates still holds at the scale of the single building. The work followed a single chain of steps: an aerial photogrammetric survey, a deep-learning classification of the resulting point cloud, a LOD2.1 reconstruction of the buildings, a roof-suitability analysis that combines solar irradiation with a supervised classifier, and finally an hourly self-consumption analysis of the residential buildings.

The main contributions can be summarised as follows.

First, the *scale of the input data.* Whereas existing city-scale assessments work on a digital surface model of about 1 m resolution (Usta et al., 2025a, 2025b), this work starts from a photogrammetric survey and builds a much finer, decimetre-level surface model (a ground sampling distance of 7.57 cm and a roof analysis carried out at 5 cm). The good resolution and accuracy of this starting data — placed in absolute coordinates to roughly 0.5 m against independent check points — are precisely what make a genuine building-by-building analysis possible: individual roof faces, dormers and obstructions that a 1 m model cannot resolve become visible and measurable. This is the geomatic foundation of the whole work, and the reason a photogrammetric, laboratory-based approach was adopted; it connects to the built-heritage and 3D city-model modelling developed in the geomatics laboratory (Aboohamzeh, Avena & Spanò, 2024).

Second, the *unbundling of the fixed area coefficient.* The rooftop-PV area normally estimated as a single fraction of the building footprint was broken into a sequence of separately measurable quantities — gross roof area, available area, and classifier-suitable area — each computed at the resolution at which it is meaningful and propagated per building rather than as a city aggregate. This makes it possible to see not only how much the coefficient over- or under-estimates on average, but where and why it fails.

Third, the *per-building result.* Across the study area the coefficient over-estimates the truly suitable roof area by about 44.5 % on average, and building by building it is unreliable, ranging from roughly 0.70 to 2.42 times the measured area. The apparent agreement at the aggregate scale therefore conceals large per-building errors — a finding with direct relevance for incentive allocation and heritage decisions, which are taken at the level of the individual building.

Fourth, the *methods and tools developed for this work*: a deep-learning point-cloud classification and a roof-suitability classifier, both trained on data labelled specifically for this thesis, together with an open, reproducible pipeline and an interactive 3D viewer for exploring the results.

The application to a dense historic centre is itself a contribution. The coefficient-based methods normally used are calibrated on modern cities and on the more regular building stock of the periphery, and they do not transfer to a dense historic centre, with its complex roof geometry, intercardinal orientations and many small obstructions — a fabric that large-scale studies usually overlook. The self-consumption analysis shows that such a centre is PV-undersized: for a typical residential building most of the production is consumed on site, yet rooftop PV can supply only about a quarter to a third of the annual electricity demand, because the limited and often intercardinally-oriented roof space cannot be matched to the evening-peaked domestic demand.

The work also has limitations that point to future development. Some inputs — the availability coefficient, the demand profile and the performance parameters — rely on standard or literature values that would benefit from local calibration; the analysis covers one portion of one historic centre and should be tested on further areas; and storage and energy-community sharing, which could raise the self-consumption of an undersized centre, were not modelled. Completing these extensions, and applying the same building-by-building method to other historic centres, are natural next steps.

References

Aboohamzeh, A., Avena, M., & Spanò, A. (2024). Automating built heritage modelling for the integration into 3D city models. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-4/W10-2024, 5–12. https://doi.org/10.5194/isprs-archives-XLVIII-4-W10-2024-5-2024

Alfio, V.S., Pepe, M., & Costantino, D. (2024). The use of Random Forest for the classification of point clouds in the urban scene. *International Journal of Engineering Trends and Technology*, 72(3).

ARERA — Autorità di Regolazione per Energia Reti e Ambiente (2023). *Prelievo medio orario per tipologia di cliente e consumi tipo dei clienti domestici.* ARERA reporting portal. https://reporting.arera.it/SASVisualAnalytics/?reportUri=%2Freports%2Freports%2Fff58199d-f010-4 sso_guest=true&reportViewOnly=true&reportContextBar=false&sas-welcome=false (accessed 2025). (Hourly withdrawal profiles by customer type; source of the residential demand shape and the *cliente tipo* reference consumption.)

Bergamasco, L., & Asinari, P. (2011a). Scalable methodology for the photovoltaic solar energy potential assessment based on available roof surface area: application to Piedmont Region (Italy). *Solar Energy*, 85(5), 1041–1055. https://doi.org/10.1016/j.solener.2011.02.022

Bergamasco, L., & Asinari, P. (2011b). Scalable methodology for the photovoltaic solar energy potential assessment based on available roof surface area: further improvements by ortho-image analysis and application to Turin (Italy). *Solar Energy*, 85(11), 2741–2756. https://doi.org/10.1016/j.solener.2011.08.010

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., & Çöltekin, A. (2015a). Applications of 3D City Models: State of the Art Review. *ISPRS International Journal of Geo-Information*, 4(4), 2842–2889. https://doi.org/10.3390/ijgi4042842

Biljecki, F., Heuvelink, G.B.M., Ledoux, H., & Stoter, J. (2015b). Propagation of positional error in 3D GIS: estimation of the solar irradiation of building roofs. *International Journal of Geographical Information Science*, 29(12), 2269–2294. https://doi.org/10.1080/13658816.2015.1073292

Biljecki, F., Ledoux, H., & Stoter, J. (2016). An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59, 25–37. https://doi.org/10.1016/j.compenvurbsys.2016.04.005

Can, G., et al. (2021). Semantic segmentation on Swiss3DCities: A benchmark study on aerial photogrammetric 3D point cloud dataset. *Pattern Recognition Letters*, 150, 108–114. https://doi.org/10.1016/j.patrec.2021.06.004

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Chiabrando, F., Costamagna, E., & Spanò, A. (2013). Passive optical sensors and related image matching methods for 3D modelling. *Bollettino SIFET*, (2), 55–76.

Choy, C., Gwak, J., & Savarese, S. (2019). 4D spatio-temporal ConvNets: Minkowski convolutional neural networks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. https://arxiv.org/abs/1904.08755

Freitas, S., Catita, C., Redweik, P., & Brito, M.C. (2015). Modelling solar potential in the urban environment: State-of-the-art review. *Renewable and Sustainable Energy Reviews*, 41, 915–931. https://doi.org/10.1016/j.rser.2014.08.060

Graham, B., Engelcke, M., & van der Maaten, L. (2018). 3D semantic segmentation with submanifold sparse convolutional networks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

ISTAT — Istituto Nazionale di Statistica (2021). *Censimento permanente della popolazione e delle abitazioni*. Roma: ISTAT. https://www.istat.it/. (Resident-family counts per census block.)

Kraus, K. (2007). *Photogrammetry: Geometry from Images and Laser Scans* (2nd ed.). Berlin: Walter de Gruyter.

Kutzner, T., Chaturvedi, K., & Kolbe, T.H. (2020). CityGML 3.0: New functions open up new applications. *PFG — Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88(1), 43–61. https://doi.org/10.1007/s41064-020-00095-z

Ledoux, H., et al. (2019). CityJSON: A compact and easy-to-use encoding of the CityGML data model. *Open Geospatial Data, Software and Standards*, 4(4). https://doi.org/10.1186/s40965-019-0064-0

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *IEEE International Conference on Computer Vision (ICCV)*, 2980–2988. https://arxiv.org/abs/1708.02002

Luthander, R., Widén, J., Nilsson, D., & Palm, J. (2015). Photovoltaic self-consumption in buildings: A review. *Applied Energy*, 142, 80–94. https://doi.org/10.1016/j.apenergy.2014.12.028

Meteonorm (2023). *Meteonorm: global meteorological database* (Version 8). Bern: Meteotest AG. https://meteonorm.com. (Linke turbidity input for the r.sun model.)

Mutani, G., & Todeschi, V. (2021). Optimization of costs and self-sufficiency for roof integrated photovoltaic technologies on residential buildings. *Energies*, 14(13), 4018. https://doi.org/10.3390/en14134018

Mutani, G., Alehasin, M., Usta, Y., Fiermonte, F., & Mariano, A. (2023). Statistical building energy model from data collection, place-based assessment to sustainable scenarios for the city of Milan. *Sustainability*, 15(20), 14921. https://doi.org/10.3390/su152014921

Peters, R., Dukai, B., Vitalis, S., van Liempt, J., & Stoter, J. (2022). Automated 3D reconstruction of LoD2 and LoD1 models for all 10 million buildings of the Netherlands. *Photogrammetric Engineering & Remote Sensing* (preprint arXiv:2201.01191). https://arxiv.org/abs/2201.01191

PVGIS (2023). *Photovoltaic Geographical Information System*. European Commission, Joint Research Centre. https://re.jrc.ec.europa.eu/pvg_tools/. (Diffuse and global radiation and typical meteorological year.)

Regione Piemonte (2025). *Punti Fiduciali* [geographic dataset]. Geoportale Piemonte. https://www.geoportale.piemonte.it/geonetwork/srv/eng/catalog.search#/metadata/c_l219:80143db5-3d23-45ca-8a91-10c26dd29ae4 (accessed 2025).

Reinhart, C.F., & Cerezo Davila, C. (2016). Urban building energy modeling — A review of a nascent field. *Building and Environment*, 97, 196–202. https://doi.org/10.1016/j.buildenv.2015.12.001

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Schönberger, J.L., & Frahm, J.-M. (2016). Structure-from-Motion revisited. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* https://doi.org/10.1109/CVPR.2016.445

Šúri, M., & Hofierka, J. (2004). A new GIS-based solar radiation model and its application to photovoltaic assessments. *Transactions in GIS*, 8(2), 175–190. https://doi.org/10.1111/j.1467-9671.2004.00174.x

Tang, H., et al. (2022). TorchSparse: Efficient point cloud inference engine. *Proceedings of Machine Learning and Systems (MLSys).*

Tang, H., et al. (2023). TorchSparse++: Efficient training and inference framework for sparse convolution on GPUs. *IEEE/ACM International Symposium on Microarchitecture (MICRO)* (arXiv:2311.12862). https://doi.org/10.1145/3613424.3614303

Todeschi, V., et al. (2021). Towards energy self-consumption and self-sufficiency in urban energy communities. *International Journal of Heat and Technology*, 39(1), 1–11. https://doi.org/10.18280/ijht.390101

Usta, Y., Montazeri, A., & Mutani, G. (2025a). Feasibility analysis of integrating solar thermal technologies into district heating network with urban building energy modeling. *Energy & Buildings*, 338, 115661. https://doi.org/10.1016/j.enbuild.2025.115661

Usta, Y., Montazeri, A., & Mutani, G. (2025b). Urban building energy modeling for self-sufficiency scenarios with solar technologies in Turin, Italy. *Journal of Physics: Conference Series*, 2940, 012024. https://doi.org/10.1088/1742-6596/2940/1/012024

Vitalis, S., Labetski, A., Boersma, F., Dahle, F., Kumar, K., Stoter, J., & Ledoux, H. (2020). CityJSON + web = ninja. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, VI-4/W1-2020, 167–173. https://doi.org/10.5194/isprs-annals-VI-4-W1-2020-167-2020

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. https://doi.org/10.1038/sdata.2016.18

Winkler, C., Dabrock, K., Kapustyan, S., et al. (2025). High-resolution rooftop photovoltaic potential assessment for a resilient energy system in Ukraine. *Energy Conversion and Management: X*, 28, 101242. https://doi.org/10.1016/j.ecmx.2025.101242

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., & Yan, G. (2016). An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, 8(6), 501. https://doi.org/10.3390/rs8060501