

Satellite Stereo Data Comprehensive Benchmark for DSM Extraction

G. Mansueto[1(✉)], P. Boccardo[2(✉)], and A. Ajmar[2(✉)]

[1] ITHACA, Via Pier Carlo Boggio 61, 10138 Turin, Italy
giuseppe.mansueto@ithaca.polito.it
[2] DIST, Politecnico di Torino, Viale Mattioli 39, 10125 Turin, Italy
{piero.boccardo,andrea.ajmar}@polito.it

Abstract. The quite recent availability of satellite stereo pairs allows users to extract three-dimensional data that can be used in different domain of applications, such as urban planning, energy, emergency management, etc. This research paper aims to extract digital surface models (DSM) from satellite stereo pairs acquired by three different satellites (Deimos-2, Pléiades-1 and WorldView-3) over the area of the city of Turin (Italy). The results are then assessed in term of geometric accuracy comparing them with a cadastral point height dataset, used as benchmark. The comparison, in terms of difference height values (between the DSM and the benchmark), is calculated on a set of sample points. Just two of the generated DSM guaranteed a high height accuracy level useful for the domain of application, such as existing cartography update, emergency management, building damage assessment, roof slope and solar incoming radiation assessment. Further developments will investigate different blending techniques and software that could provide more accurate results.

Keywords: Satellite stereo pair · DEM/DSM extraction · Accuracy assessment

1 Introduction

The development of the very high resolution (VHR) optical sensors mounted on satellites nowadays allows to achieve a ground sample distance lower than 1 m. This increased, in the last years, the geo-referencing accuracy and the mapping capability, bringing to an improvement in resolution of digital surface models (DSMs) based on data acquired by VHR sensors [1]. Since these data models permit more precise and economically advantageous measurements of large land surfaces, they are more and more required in many different application domains and new methodologies have been implemented to improve the quality and accuracy of automatic 3D processing [2]. The capacity of acquiring high-resolution stereo couples, from which to extract accurate 3D data, is therefore considered a relevant element in new satellites.

The aim of the paper is to address the differences in resolution and accuracy of the generated DSM over the city of Turin (Italy) from a satellite stereo pair acquired by three different satellite sensors, comparing the result in term of geometric accuracy using two benchmark, a ground height points dataset of the City of Turin, constituted by geodetic heights values of ground marked points (mainstay, trustworthy, network

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12252, pp. 858–873, 2020.
https://doi.org/10.1007/978-3-030-58811-3_61

vertex and stable reference points) and a cadastral gutter height points dataset, reporting the geodetic gutter heights of each building of Turin. The paper describes the common datasets and procedures for the DSM generation, exploring the potential of stereoscopic images of Deimos-2, Pléiades-1 (in two- and tri-stereo mode) and WorldView-3 satellites for this purpose and assessing them from qualitative and quantitative point of view.

The DSM is generated by collecting tie points (TPs) or ground control points (GCPs) from a stereo pair, then used to create the Epipolar images from which the automatic DSM generation is carried out.

The paper is organized as follows. After this introduction, Sect. 2 provides a description of the satellites platforms, on-board sensors and acquisition modes. Section 3 introduce the stereo imagery used dataset. In Sect. 4 it is described the processing methodology adopted to generate the satellite-based DSMs, then qualitatively analyzed in Sect. 5. Section 6 focuses on the description of benchmark data characteristics, used for accuracy verification. Section 7 assess the satellite-based DSM accuracy, discussing about the results and, in the end, chapter 8 close with conclusion and future developments.

2 Satellite and Sensor Description

The input dataset is composed by images acquired by civilian VHR stereo satellites sensors (Table 1), which allows creating first-rate quality stereo geometry thanks to a convergence angle higher than 0.5, improving the DSM vertical accuracy.

Table 1. Stereo satellites comparative table

Satellite	GSD [m] (Panchromatic)	Stereo	Tri-stereo
Deimos-2	1	y	n
WorldView-3	0.31	y	n
Pléiades-1	0.5	y	y

This sun-synchronous satellites (Deimos-2, WorldView-3 and Pléiades-1) can acquire more than one image, with a difference inclination of the platform, in a low-range acquisition time difference (1–3 min), minimizing the radiometric variation and facilitating the DSM generation [3].

2.1 Deimos-2

With a lifetime of at least seven years, Deimos-2 is a VHR stereo multispectral optical satellite (1 m GSD resolution at nadir in panchromatic), equipped with a push-broom VHR camera with 5 spectral channels (1 panchromatic, 4 multispectral), improved by the fast and precise rotation of the platform which hold it [4].

2.2 WorldView-3

WorldView-3, launched in 2016, is the latest in a constellation of commercial high-spatial resolution Earth imaging satellites developed by Maxar (DigitalGlobe Inc). It is equipped with a VHR camera with 29 spectral channel: 1 pan-chromatic (with 0.31 m GSD), 8 multispectral, 8 SWIR bands (from 1.2 to 3.7 m GSD resolution), 12 CAVIS bands ("clouds, aerosols, water vapor, ice, and snow" at 30 m GSD resolution for atmospheric compensation), improved by the fast and precise rotation of the platform which hold it [1].

2.3 Pléiades-1

Pléiades-1 is a very-high resolution stereo multispectral optical satellite constellation (0.5 m GSD resolution in panchromatic at nadir), composed by two spacecraft purchased by CNES (Space Agency of France) company. The satellite is equipped with a VHR camera with four spectral bands (blue, green, red, and IR), capable of acquiring high-resolution stereo imagery in just one pass and accommodating large areas (up to 1,000 km × 1,000 km) [5]. One of the aims of this mission is the provision of so-called "level-2 products" to customers, consisting of a panchromatic image with a merged multispectral image orthorectified on a digital terrain model (DTM) [6].

Pléiades-1 offers also a tri-stereo terrain data generation approach (Fig. 1), which differs from conventional stereo data generation through the application of two oblique and one near-nadir viewing of the terrain, as opposed to just two oblique views, which provides the ideal solution for accurate 3D modelling. This is especially relevant in areas of high relief variation, including dense, high-rise urban landscapes, where the tri-stereo image coverage image combination significantly minimizes the problem of data 'loss error' areas in the final DSM. This can result in a reduction of ±75% of "hidden area objects", which can arise with conventional 2 × oblique image coverages, as a result of object lean and view obscuring effects [7].

3 Dataset

The input dataset is composed by seven VHR stereo satellite images, two from Deimos-2, two from Worldview-3 and a triplet from the tri-stereo satellite Pléialdes-1. All the used panchromatic images have been provided with a low processing data level, with basic geometric and radiometric correction. As we can see in the Tables 2 and 3, the two/three stereoscopic images coming from the same satellite are acquired on the same date, with a difference of few minutes (the maximum difference is 9'47") in acquisition time, in order to minimize solar irradiation differences that may cause shadows-linked errors.

The chosen test area is the city of Turin, regional capital of Piedmont in Italy. Figure 2 outlines the stereo pairs orientation in the space, giving an idea of their dimension, and the chosen test area for the further accuracy assessment.

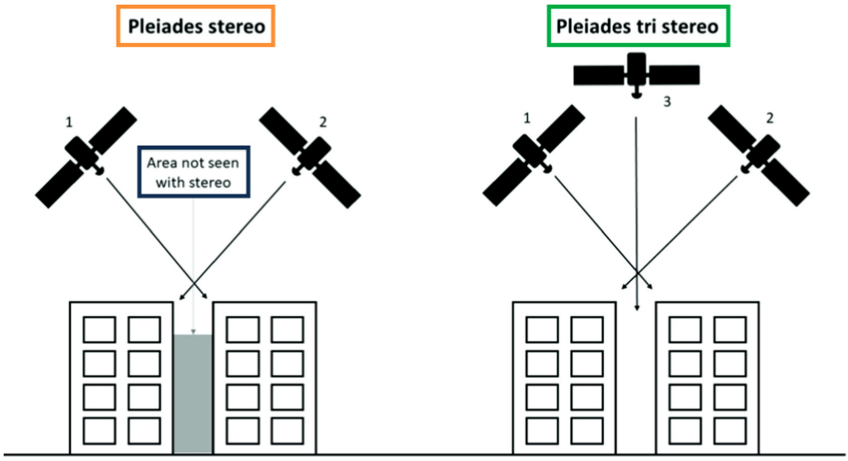

Fig. 1. Comparison between stereo and tri-stereo Pleiades acquisition [8]

Table 2. Characteristics of Deimos-2 (D2) and WorldView-3 (W3) panchromatic images

Image ID	D2 (1)	D2 (2)	W3 (1)	W3 (2)
Acquisition date	11/07/18	11/07/18	16/12/17	16/12/17
Acquisition time (GTM)	10:04:50	10:06:34	11:05:21	11:06:16
Incidence angle (DEG)	4	1.4	−25.0	−26.1
Sun azimuth (DEG)	133.93	134.68	175	175.2
Sun elevation (DEG)	60.60	60.85	21.6	21.6
Columns	11712		33333	
Rows	8604		33333	
Framed area (km^2)	196.3		99.4	

Table 3. Characteristics of Pléiades-1 (PH1) panchromatic images

Image ID	PH1 (1)	PH1 (2)	PH1 (3)
Acquisition date	27/04/2018		
Acquisition time (GTM)	13:23:08	13:24:06	13:32:55
Incidence angle (DEG)	−4.69	−5.93	−4.59
Sun azimuth (DEG)	153.99	153.99	153.99
Sun elevation (DEG)	56.02	56.02	56.02
Columns	21340	21482	21887
Rows	21296	22356	22736
Framed area (km^2)	128.3		

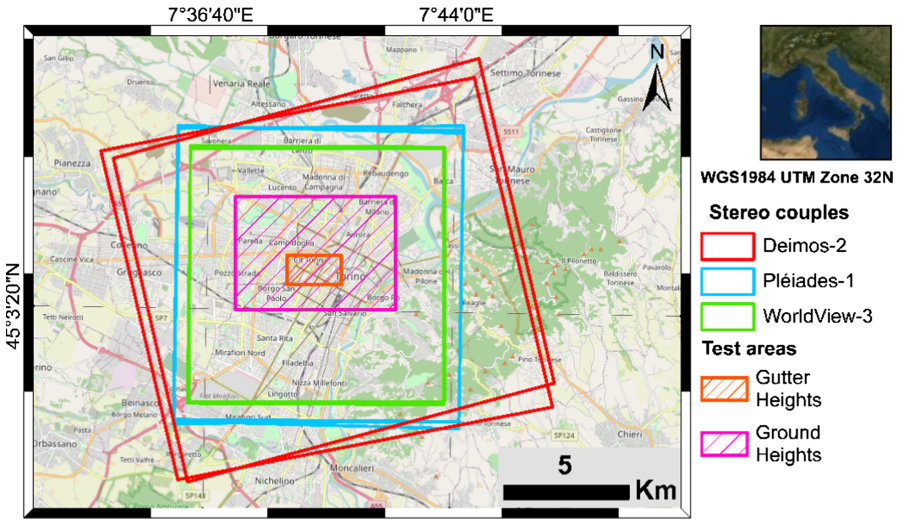

Fig. 2. Location of satellite imagery footprints and test areas in Turin, OSM Basemap in background

4 Processing Methodology

PCI Geomatica, and specifically the OrthoEngine production toolkit, has been used for stereo satellite data processing. Using this software, it is possible to generate a DSM referred to the ellipsoid height or to the geodetic elevation [9].

The steps followed in the extraction process are:

1. Creation of a new project choosing the current satellite origin of the dataset and as orientation mode the one with RPCs model (exploiting the metadata file supplied by the satellite data provider);
2. Import of the stereo pair and conversion in.pix file (PCI Geomatics internal file format);
3. Tie points (TP) collection by means of automatic extraction, imposing the following settings:
   a. Distribution Pattern: Overlap area
   b. Trial per point (number of iterations per point): 3
   c. TP per area: 50
   d. Min. acceptance score: 0.75
   e. Search radius: 100 pixels
   f. Sample source method: Susan
   g. Elevation Search Strategy: SRTM DEM with 30 m GSD extracted over the area

The residuals RMSE (average between the RMSE on the x and y axis) of each TP collected needs to be a tenth of the dimension of the sampling scheme (pixels size of the image) "to ensure that they will be completely independent of their random position and their orientation in relation to the sampling scheme (Shannon Sampling theorem)" [8].

4. Epipolar image generation
5. Automatic DEM extraction, inserting the following inputs as settings:
   a. Extraction method: Normalized cross-correlation (NCC)
   b. Elevation range: Automatic
   c. DEM detail: High
   d. Terrain type: Hilly (in this case)
   e. Output DEM vertical datum: Mean sea level
   f. Output DEM channel type: 32-bit real
   g. Pixel sampling interval 2
   h. Smoothing filter: Low
   i. Geocoded DEM resolution: double of sensor GSD (i.e. 1 m for Pléiades-1)
   j. Output option: Blending

It is important to point out that, due to the PCI Geomatics processing, the resolution of the generated DSMs is always the half of the one in the used dataset [10].

5 Qualitative Analysis

There were generated four models: one Deimos-2, one WorldView-3, one Pléiades-1 two-stereo and one Pléiades-1 tri-stereo DSM.

From a preliminary and qualitative assessment, evaluating noise, sharpness in city canyons, ability in shaping the geometries and identification of details, emerged that Deimos-2 model (with a resolution of 2 m) (Fig. 3) results very poor with respect to the others (Fig. 4 Pléiades-1 two-stereo DSM for comparison), showing an elevation accuracy that does not permit to identify or distinguish the shape of the object on the ground and presenting several issues related to noise and false elevation spikes.

Concerning the other DSM, it was found that for WorldView-3, even if presents a high GSD in the acquisition, the generated DSM quality results lower in comparison to Pléiades-1 ones. Focusing on the roof in the Fig. 5, reported as example, its shape is not represented faithfully in the WorldView-3 model. Both Pléiades-1 DSMs shows a good result; the two-stereo DSM better represents the roofs flaps, exhibiting that the external spans of it are higher than the central one. The tri-stereo one, instead, better defines the details in the North-West part.

As far as the streets and city-canyon are concerned (Fig. 6), tri-stereo Pléiades-1 DSM reach up a higher quantity of details in the scene, outlining more objects in the middle of main street. The two-stereo Pléiades DSM, instead, delineate the streets shaping them completely, without the noise founded in the other one and without the building overlaps present in WorldView-3 DSM, which also in this case looks weak in comparison to the others.

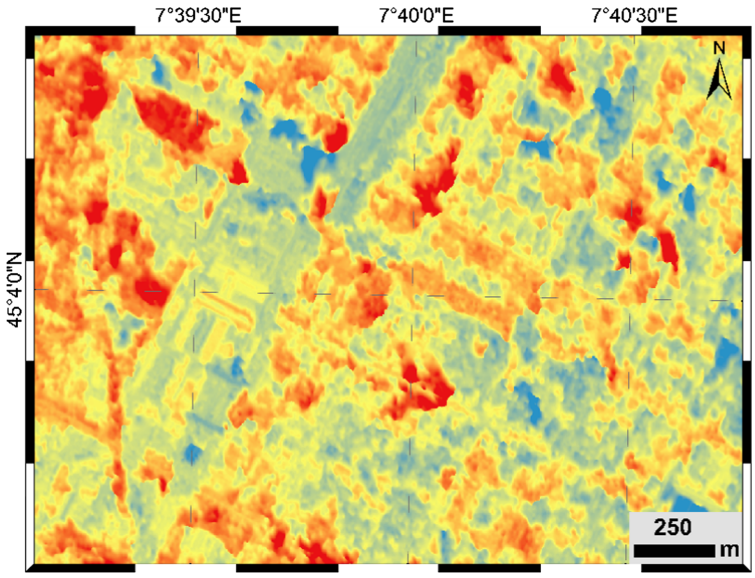

Fig. 3. Focus of Deimos-2 geoid DSM in the city area of Turin

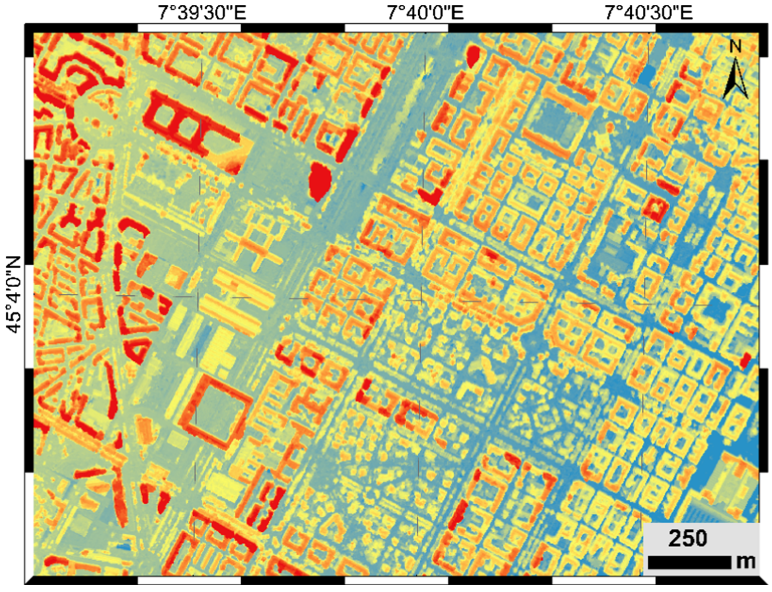

Fig. 4. Focus of Pléiadel-1 two-stereo geoid DSM in the city area of Turin

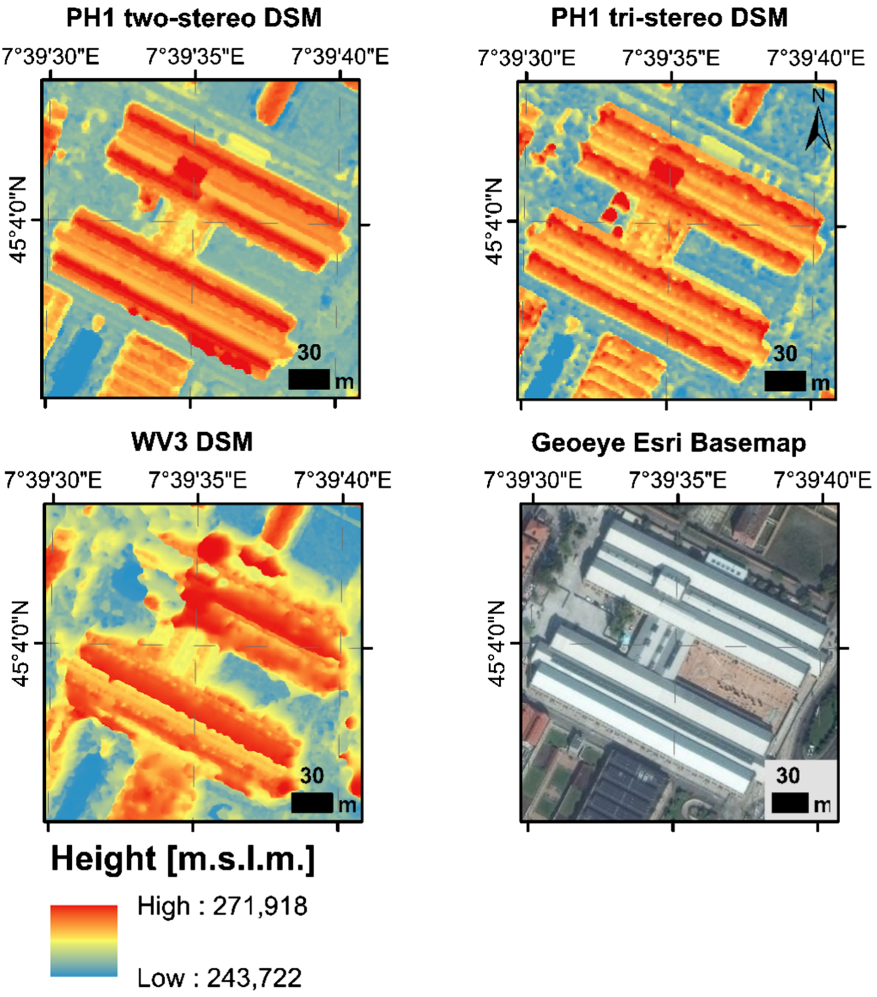

Fig. 5. Visual analysis over a limited urban area, in order: Pléiades-1 two-stereo, Plèiades-1 tri-stereo, WV-3 extracted geoid DSMs and Geoeye Esri basemap

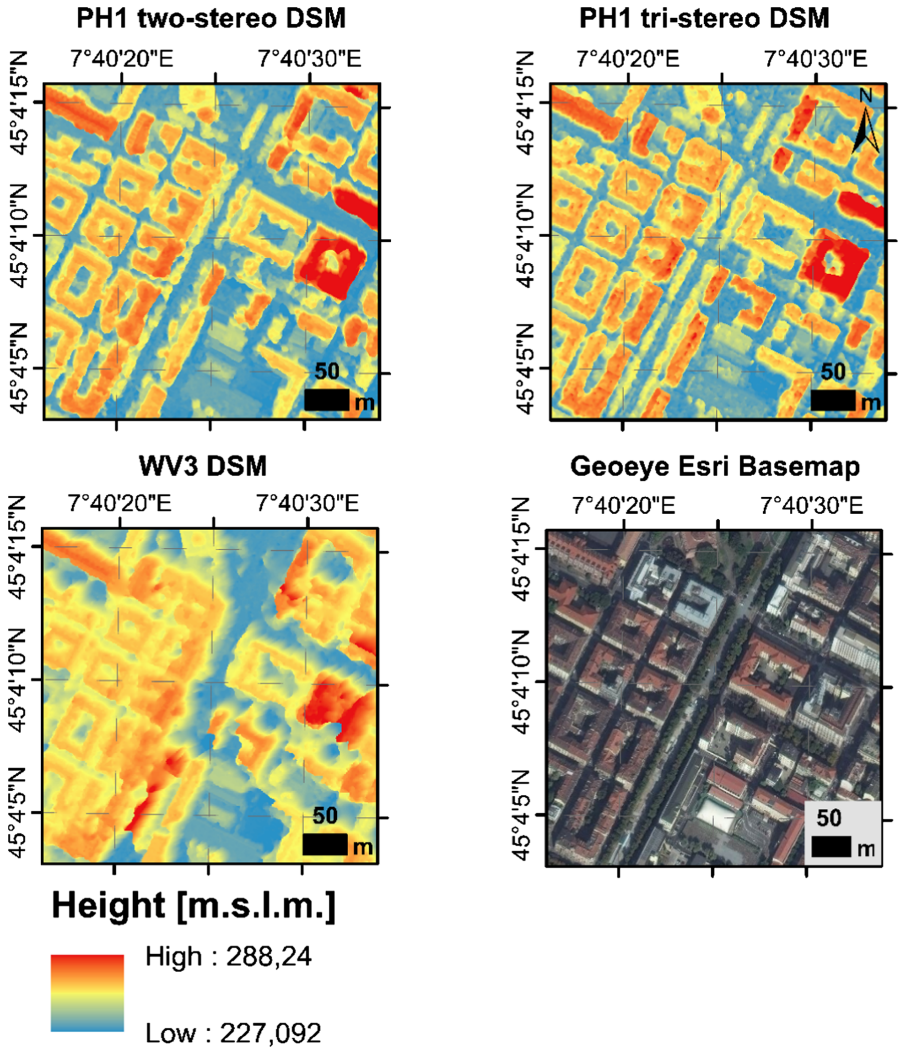

Fig. 6. Visual analysis over a limited urban area, in order: Pléiades-1 two-stereo, Plèiades-1 tri-stereo, WV-3 extracted geodetic height DSMs and Geoeye Esri basemap

6 Description of the Used Benchmark Data

The qualitative output was double checked with the quantitative analysis of the data, evaluating the altimetric accuracy of the models through the use of two benchmarks documenting the correct geoid heights of Turin: a ground height points dataset (last update 2011) of the City of Turin, constituted by geodetic heights values of a sample of ground marked points (mainstay, trustworthy, network vertex and stable reference points) and a cadastral gutter height points dataset (last update 2015), reporting the geodetic gutter heights of each building of Turin.

In order to compare them with the DSMs, these benchmarks have been converted into two raster products:

1. a raster reporting in the pixel cells the values of ground heights ("ground raster");
2. a raster reporting in the pixel cells the values of gutter heights of the buildings ("gutter raster").

The comparison was carried out in two smaller test areas, common for all the satellite acquisition frames, respectively of 27,6 km^2 for ground and 2,4 km^2 for gutter heights assessment.

At this point, the benchmark have been subtracted to the six assessed DSMs (2 DSMs for each of the three sensors used as input), obtaining three ground height difference rasters and three gutter height differences raster (as output), and then (after the subtraction) re-converted in vector products (in order to extract the statistics). The achieved vector points dataset, composed by three ground height differences and three gutter height differences, describes the accuracy in "geodetic height difference" between the DSMs data and a certain-known one.

7 Results and Discussion

As said previously, the DSMs assessment (WV3, PH1 two- and tri-stereo) has been carried out on three ground height differences and three gutter height differences. In the Fig. 7 it is reported as example the test area within the "ground height difference" points form the PH1 tri-stereo subtraction, against the back-ground of PH1 tri-stereo DSM.

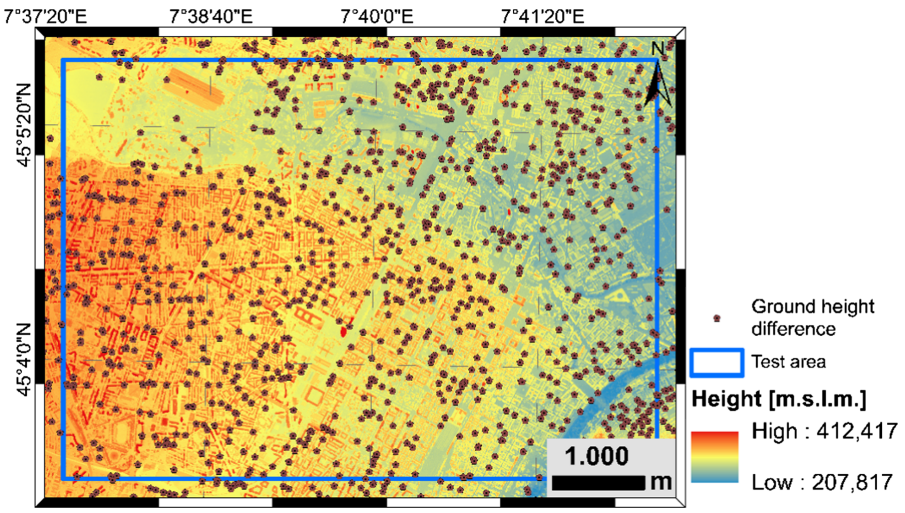

Fig. 7. Test area, inside "ground height difference" points PH1 tri-stereo referred, PH1 tri-stereo DSM on the background

From these points we have extracted the statistics for all the six outputs (count in the sample, minimum, maximum, mean and standard deviation above all samples heights). All the values are reported in the Table 4 and 6.

7.1 Results Using Ground Heights Benchmarks

The analysis on ground height differences was conduct on a sample of 2693 points for WorldView-3 and 988 points for Pléiades-1 (each representing a height difference).

Table 4. Summary table of the statistics referred to the ground height differences

GROUND HEIGHTS	WV-3	PH1 2-st	PH1 3-st
Count	2693	988	988
Min [m]	−104	−7	−21
Max [m]	50	56	47
Mean [m]	5.7	11.8	1.5
St. dev [m]	9.4	9.4	10.2

Being a difference between the heights of the DSM and the real heights, the models reach the perfection for values of difference that tends to zero.

Looking at the statistics referred to all the sample (Table 4), focusing on the standard deviation (expressing the dispersion around the mean), it is possible to notice that it is high for all the generated DSMs, meaning that outliners are present in the analyzed sample. Moreover, looking at the mean, it is visible that all the models seems to overestimate (high values of mean) the city height values, meaning that the most of the outliners present high values of heights (spoiling the statistics on the model).

The anomalies that led to these phenomena can be related to two cases:

a. High positive difference values (Fig. 8), found where the point (which in the reality is on the ground) fall in the DSM on a roof, then the DSM present the height of the roof and the point the ground one, and consequently a high height difference value; the effect is related to a perspective geometric distortions, due to the off nadir angle in the satellite acquisition.

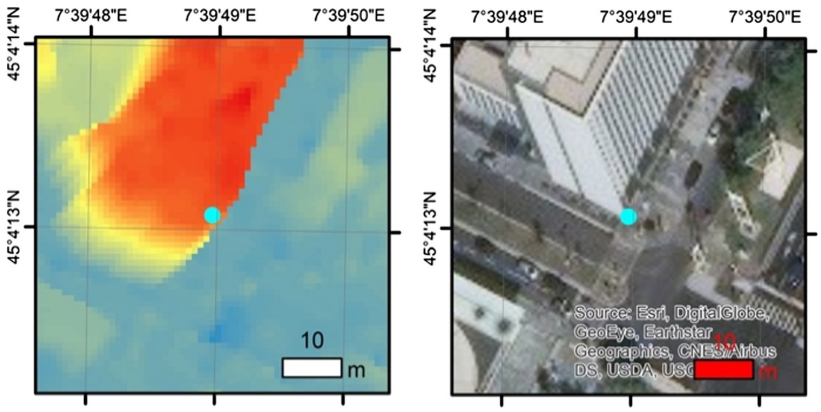

Fig. 8. Example of an "High positive difference values anomaly" in ground heights assessment (left PH1 two-stereo DSM, right Geoeye Esri basemap)

b. High negative difference values, when the DSM for the foreshortening or the presence of shadows (usually in the city canyons) has wrong digitalized the ground, underestimating the height value. In this case the DSM present a really low height and the point the real one of the ground, implying a negative height difference value.

It is observable that PH1 tri-stereo declare a low value of mean, really distant from the high standard deviation one. This is amenable to the fact that this sensor catch more details on the ground, bringing to another additional effect, that are errors caused by the presence of object on the streets (e.g. vehicles, dehors, stalls, etc.) during the acquisition of the images, bringing in this cases to an overestimation of the heights.

After this ascertainment, it was decided to restrict the sample in an acceptance range between −10 m and 10 m. In this way, deleting the systematic errors from the sample, it was obtained a more representative one.

From a first look at the new samples statistics (Table 5), we can assert that no one of the three models can be evaluated as the best one.

Table 5. Summary table of the statistics referred to the ground height differences in the range [−10, 10]

GROUND HEIGHTS	WV-3	PH1 2-st	PH1A 3-st
Points in [−10; 10]	64%	54%	76%
Mean [m]	0.3	4.3	−2.9
St. dev [m]	5	2.8	5.4

The model with the better distribution of the values in the range, presenting a Gaussian trend, is PH1 tri-stereo. It is also the model with mayor number of sample points falling in the acceptance range (76%), followed by WV3 with 64% and PH1 two-stereo with 54%. As it is possible to see the mean of the tri-stereo DSM is

negative, indicating that the "high positive difference" issue is minimized with respect to the two-stereo one, due to the tri-stereo mode which reduce the perspective geometric distortions. Despite this, this model displays a high value of standard deviation, with very disperse values far from the mean. In WV3, and even more in Pléiades two-stereo, the distribution is concentrated on positive values, suggesting that these models reduce the fore-shortening effect. Discussing about these two models: even if the mean of WV3 DSM ground heights is proximal to zero the values of the samples are not approaching to the fitting one, PH1 two-stereo DSM instead display a lower value of standard deviation at the expense of a higher mean. Therefore, the analysis shows a better ground heights accuracy in PH1 two-stereo model.

7.2 Results Using Gutter Heights Benchmarks

In this case the sample is made by a really high number of points (262690 for WV3, 91835 for PH1) displaced on the gutters (perimeter walls) of the buildings. All the three models exhibit high values of standard deviation (Table 6). The mean, instead, is oriented to negative heights for WV-3 and PH1 in tri-stereo mode and to positive for the two-stereo model of PH1.

Table 6. Summary table of the statistics referred to the gutter height differences

GUTTER HEIGHTS	WV-3	PH1 2-st	PH1 3-st
Count	262690	91835	91835
Min [m]	−74	−62.7	−54
Max [m]	90	69.3	61.9
Mean [m]	−2.3	3.1	−6.6
St. dev [m]	7.5	6.2	7.7

Also, in this case the models are affected by issues related to a perspective geometric distortion, due to the off-nadir angle in the satellite acquisition:

a. High negative difference values, found where a point of the benchmark (which in the reality is on the gutter) falls in the DSM on the ground area, then the DSM present the height of the ground and the point the gutter one (Fig. 9), and consequently a negative height difference value;

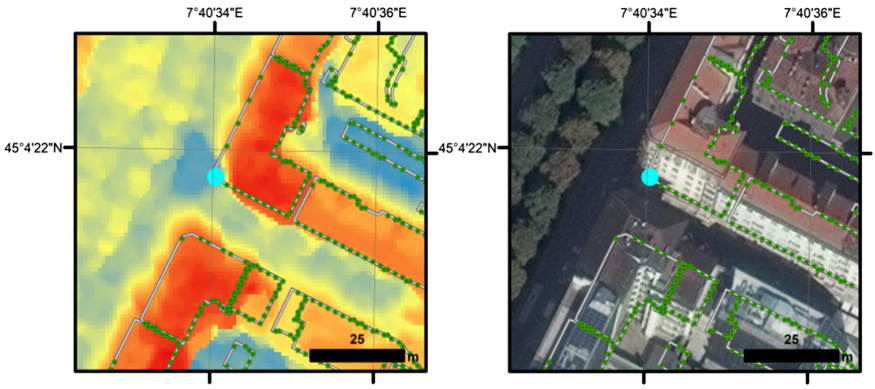

Fig. 9. Example of an "High negative difference values anomaly" in gutter heights assessment (left PH1 two-stereo DSM, right Geoeye Esri basemap)

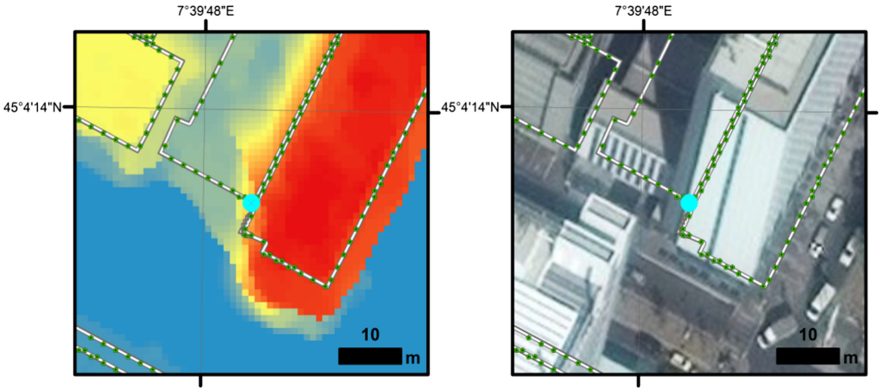

Fig. 10. Example of an "High positive difference values anomaly" in gutter heights assessment (left PH1 two-stereo DSM, right Geoeye Esri basemap)

b. High positive difference values, when the DSM has overestimated the height value. In this case the error can also be related, as it is possible to appreciate in the Fig. 10, to strong heights differences between adjacent buildings or to the presence of new constructions in the DSMs (WV3 2017, PH1 2018) not present in the benchmark dataset (2015).

Although the sample is numerous, in order to obtain a more representative sample delating the systematic errors, also in this casa it was preferred to reduce the acceptance range from −10 to 10.

All the models frequency distributions present a gaussian trend (Fig. 11). It is important to point out that the higher pecks of WV3 distribution is due to the fact that the sample is denser of points (because of the better resolution), falling in the current case in that range of values. The different position of the distribution peaks of PH1 two-

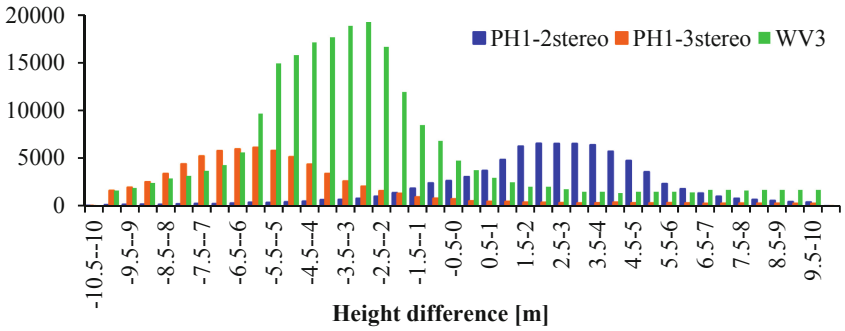

Fig. 11. Frequency distributions of gutter height differences of the three satellites in the range [−10,10]

Table 7. Summary table of the statistics referred to the gutter height differences in the range [−10, 10]

GUTTER HEIGHTS	WV3	PH1 2-st	PH1 3-st
Points in [−10; 10]	85.2%	89%	78.9%
Mean [m]	−2.5	2.3	−4.8
St. dev [m]	3.8	3	3.7

and tri- stereo DSMs confirm the previous statements. Looking at the percentage of sample points falling in this range (Table 7), it is possible to conclude that in the city the DSMs better defines the roofs than the streets, due to the fore-shortening and inclination. Analyzing the statistics, it is, once again, evident that PH1 two-stereo model guarantee the best gutter heights accuracy. At the end of this assessment of the gutter heights differences is legitimate to evaluate PH1 two-stereo DSM as the best product, confirming the conclusion of the qualitative analyses.

8 Conclusion and Future Developments

It was established that the quality of a DSM has not a resolution dependence only. Despite to what expected, concerning the other DSM, it was found that for WorldView-3, even if presents a high GSD in the acquisition, the generated DSM has lower accuracy in comparison to Pléiades-1 one, showing in particular some problems in the canyon shaping. Pléiades-1 DSMs exceed in height accuracy level WorldView-3 product. Focusing on Pléiades-1 models, the one generated in tri-stereo mode exhibit higher capacities in the identification of details, small objects on the ground (e.g. cars, vegetation, road signs), which in some case can turn into noise; the model generated in two-stereo mode, instead, demonstrate an improved sharpness in city canyons and roofs geometries shaping. Resulting more efficient in terms of both ground and building heights calculation, it is therefore legitimate to evaluate PH1 two-stereo DSM as the best product over the generated ones.

If in the next future more stereo pairs will be available over the area of Turin, it could be very interesting to investigate on them in order to find other discriminating factors, testing them with different blending techniques and more accurate bench-marks (LIDAR) for the quantitative assessment.

References

1. Hu, F., Gao, X.M., Li, G.Y., Li, M.: DEM extraction from worldview-3 stereo-images and accuracy evaluation. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 2016, pp. 327–332 (2016)
2. Poon, J., Fraser, C.S., Chunsun, Z., Zhang, L.I., Gruen, A.: Quality assessment of digital surface models generated from ikonos imagery. Photogramm. Rec. **20**(110), 162–171 (2005)
3. Aguilar, M.Á., et al.: DSM extraction from worldview.pdf. IEEE Trans. Geosci. Remote Sens. Gener. (2013)
4. Galileo, E., Imaging, S.L.U.: DEIMOS-2 Imagery User Guide (2015)
5. Airbus Defence & Space: Pléiades Spot the Detail, Geo-Intelligence Fact Sheet (2013). https://www.intelligence-airbusds.com/files/pmedia/public/r49228_9_pleiades_product.pdf
6. Codou, G., Cubero-Castan, E., Duverger, T., Mesnard, R., Tavera, F., Casserra, C.: Operational scheduling of direct tasking innovative concept to improve reactivity on earth observation system. In: SpaceOps 2016 Conference (2016)
7. Thompson, M., et al.: High resolution landcover modelling with pléiades imagery and DEM data in support of fine scale landscape thermal modelling. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-3/W2 (2017)
8. Blaschke, T.: Object based image analysis for remote sensing (2010)
9. Boccardo, P., Sandu, C., Ajmar, A., Perez, F.: Digital surface models extraction by Göktürk-1 satellite stereo pairs. In: 9th International Conference on Recent Advances in Space Technologies RAST (2019)
10. PCI: Geomatica OrthoEngine Pleiades 1A data – DEM extraction and DSM to DTM Conversion, pp. 1–12 (2013)