remote sensing

MDPI

Technical Note

Metric Accuracy of Digital Elevation Models from WorldView-3 Stereo-Pairs in Urban Areas

Emanuele Mandanici *,†, Valentina A. Girelli † and Luca Poluzzi †

Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, 40136 Bologna, Italy; valentina.girelli@unibo.it (V.A.G.); luca.poluzzi5@unibo.it (L.P.)

* Correspondence: emanuele.mandanici@unibo.it

† Current address: Viale del Risorgimento 2, 40136 Bologna (BO), Italy.

Received: 6 March 2019; Accepted: 9 April 2019; Published: 11 April 2019

Abstract: WorldView-3 satellite is providing images with an unprecedented combination of high spatial and spectral resolution. The stereo capabilities and the very high resolution of the panchromatic band (0.31 m) have been fostering new applications in urban areas, where the complexity of the morphology requires a higher level of detail. The present technical note aims to test the accuracy of digital elevation models that can be obtained by WorldView-3 stereo-pairs in these particular contexts, with an operational state-of-the-art algorithm. Validation is performed using check points and existing models of the area (from LiDAR data and oblique aerial images). The experiments, conducted over the city of Bologna (Italy) with six images, proved that roof surfaces and open spaces can be reconstructed with an average error of 1–2 pixels, but severe discrepancies frequently occur in narrow roads and urban canyons (up to several metres in average). The level of completeness achievable with only one pair is extremely variable (ranging from 50% to 90%), due to the combined effect of the geometry of acquisition and the specific urban texture. Better results can be obtained by using more than one pair. Furthermore, smaller convergence angles can be beneficial for the reconstruction of specific urban structures, such as soaring towers.

Keywords: WorldView-3; urban modelling; DEM accuracy; DEM validation; semi-global matching

1. Introduction

WorldView-3 (WV3) satellite was launched in August 2014 and is able to provide multispectral images with an unprecedented combination of high spatial and spectral resolutions. The platform is equipped with a 0.31 m resolution panchromatic sensor, an eight-band visible and near-infrared imager (operating between 397 and 1039 nanometres) with a ground sample distance (GSD) of 1.24 m, a short-wave infrared sensor with further eight bands (1184–2373 nm) at a resolution of 3.7 m, and finally the CAVIS (Clouds, Aerosol, Vapors, Ice and Snow) sensor with twelve additional bands at 30 m for the retrieval of atmospheric properties [1]. The panchromatic and multispectral sensors have also stereo capabilities and the nominal geolocation accuracy of the images is about 5 m [2], which can be greatly improved using a few ground control points [3]. WV3 platform follows a sun-synchronous orbit at a nominal altitude of 617 km and is expected to be operational for about ten years.

These characteristics foster new applications, especially for urban environments. WV3 images have been tested in many different fields, such as mapping heterogeneous agricultural landscape [4], mineralogical and lithological mapping through spectral indices [5], urban land cover classification with deep learning models [6], bathymetry retrieval [7], among the most recent ones. It is worthwhile to mention that the strategic objectives of WV3 include also the detection of buildings and the recognition of materials.

The ability to collect stereo-pairs of panchromatic images at the resolution of 0.3 m makes WV3 suitable for the production of detailed digital elevation models (DEMs). Several authors tested the stereo capabilities of WV3 in different environments, obtaining promising results. Aguilar et al. [8] tested two algorithms for DEM extraction on different land cover types and found a standard deviation of 0.2 m on bare soils and 0.7 m on plastic greenhouses (compared to a reference LiDAR survey). Similarly, Hu et al. [9] found a standard deviation of 0.6 m in the model of a mountainous area in China, using four WV3 images.

Urban environments, however, present specific challenges, due to the complex morphology of the urban texture. Some authors [10] already noted that the accuracy and completeness of models in urban areas are lower. Rupnik et al. [11] applied a multi-view reconstruction algorithm on an urban area and reported some differences in quality indicators as a consequence of varying housing density. In this context, the geometry of acquisition appears to influence the quality of the final results [12].

The present technical note aims at evaluating the accuracy of DEMs that can be obtained from WV3 imagery, which offers the highest spatial resolution among the currently operational and commercially available satellite sensors with stereo capabilities. Differently from previous works, the present note focuses exclusively on urban areas, which require very high resolution to correctly describe their complex morphology. In addition, urban environments are generally the most problematic for the reconstruction process, thus they constitute a particularly meaningful test to assess the performances of this imagery.

The evaluation of the obtained DEMs is carried out on different urban contexts, from extremely dense urban centres, with ancient buildings of variegated shapes and narrow hollows, to more recent residential and productive areas, with larger buildings, open spaces and wider roads. In this technical note, not only the accuracy of the generated 3D model as a whole is estimated, but different components of the urban texture (e.g., roofs, carriageways, road network orientation) are analysed separately to investigate their specific contributions to the global errors. In addition, the effects of different acquisition geometries are considered. All of these aspects, indeed, may be relevant for end-users to evaluate the suitability of WV3-derived models on the basis of the specific application requirements.

The experimentation exploits a state-of-the-art DEM extraction algorithm, the semi-global matching [13], that is very popular in routinely production workflows where robust and efficient stereo-matching is required, even though it is capable of processing only one single pair at time (but this is a frequent scenario with expensive satellite imagery). This algorithm is considered very robust and able to reduce large outliers in low or non-textured areas while preserving edges and sharp object boundaries. The use of pixel-wise cost functions provides the capability to resolve fine spatial structures on the object surface [14].

The following sections analyse the performance of WV3 images over the urban area of Bologna (Italy) in a production perspective.

2. Materials

For the present experimentations, three different stereo-couples were acquired over the area of Bologna (Italy) in 2017 in a few days, the first on 14 September 2017 and the other two on 20 September (Table 1). The names used in this paper (and reported in the first column) are made of the last four digits of the time stamp included in the original file name; the identification code used in the supplier's catalogue is also reported (second column). All the images were acquired at noon and can be considered cloud-free and haze-free. They cover an area of approximately 100 km^2, encompassing the territory of the Municipality of Bologna (Figure 1). The elevation angle and azimuth of the satellite at the time of acquisition are reported in Figure 2a. All the images were pre-processed by the supplier at the Ortho-Ready level, which means they underwent basic radiometric correction, they were resampled at 0.3 m square pixels and they are provided with the coefficient of a rational polynomial model for the orthorectification process.

Table 1. List of the WorldView-3 images used for the experimentation, with date and time of acquisition, off-nadir angle and average raw ground sample distance of the scene.

Image	Catalogue ID	Date	Time (UTC)	Off-Nadir (deg)	Raw GSD (m)
2938	104001003222B500	14 Sep 2017	10:29:38	18.5	0.34
2824	1040010032519A00	14 Sep 2017	10:28:24	32.9	0.43
2526	1040010033212F00	20 Sep 2017	10:25:26	22.0	0.35
2626	10400100348FBD00	20 Sep 2017	10:26:26	31.4	0.41
2549	1040010033D35E00	20 Sep 2017	10:24:58	20.8	0.35
2458	104001003248B400	20 Sep 2017	10:25:49	30.9	0.41

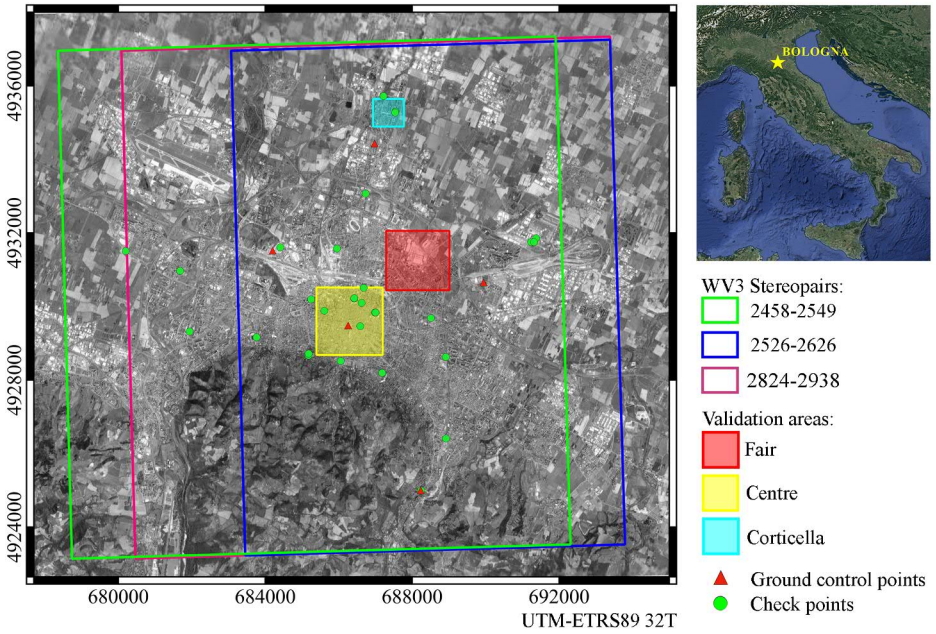


Figure 1. Location of the study area with the footprints of the WV3 pairs, the validation areas and control and check points.

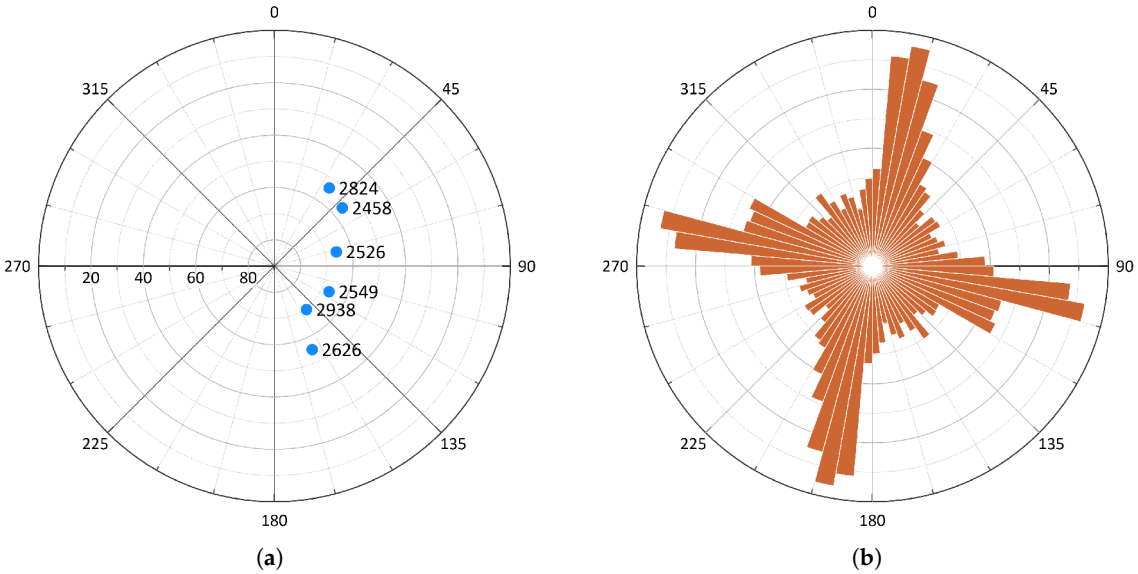


Figure 2. (**a**) sky plot with satellite azimuth and elevation angles at the time of acquisition; (**b**) polar histogram of the city centre road azimuths derived by the technical cartography.

In order to refine the rational polynomial model and to check the accuracy of the final products, a set of points were directly measured in the observed area using network real-time kinematic (NRTK) techniques, based on virtual reference stations. A dual-frequency GRS1 receiver (Topcon Positioning Systems, Livermore, CA, USA) with a PG-A1 antenna was used and connected to the NETGEO network [15]. Overall, 35 points were measured: 24 lie on the ground level, while 11 are located on flat roofs. The precision of the measured coordinates is about 2 cm for the planimetry and about 5 cm for the altimetry [16].

For a more exhaustive comparison of the generated DEMs, two existing products were collected: a point cloud obtained from a LiDAR survey performed on 26 January 2009 over the city centre and a model obtained from oblique aerial images and Structure-from-Motion algorithms over a smaller residential area.

The LiDAR survey used an ALTM 3033 (Optech, Toronto, ON, Canada) with a field of view of ± 11 degrees. Given the operational altitude of 1250 m above the ground level and the distance between strip axes of 322 m, the horizontal accuracy is 0.6 m and the vertical one 0.2 m. The LiDAR cloud has a density of 1 point/m^2 in average [17].

The model from the oblique images is described in detail in the paper by Bitelli et al. [18] and only the main characteristics are reported here. A block of 132 aerial oblique images (GSD 8 cm), acquired in May 2017 using the MIDAS5 system (Track'Air, Enschede, Netherlands), was processed by the ContextCapture Structure-from-Motion software package (version 4.4.9, Bentley Systems, Exton, PA, USA). In addition, 350 million points with an average spacing of 5 cm constitute the obtained point cloud for the area of interest, covering 1.5 km^2. The accuracy of the final model can be considered in the order of few centimetres.

Finally, the technical cartography of the Municipality of Bologna, which comes with a nominal scale of 1:2000 and a tolerance for planimetric coordinates of about 0.5 m, was used for GIS analyses. Figure 2b, for example, shows the frequency distribution of the azimuth angles of the roads in the city centre.

All the data were coherently framed in the ETRF2000 (epoch 2008), which is the official geodetic framework for Italy, using the NRTK points.

3. Methods

3.1. WorldView-3 Image Processing

All the WorldView-3 images were oriented through a rational polynomial model, using the coefficient provided by the supplier. The model was further refined with a first-order polynomial whose coefficients were estimated using five control points and 73 tie-points, well distributed in the study area and manually collimated. In all the six images, exactly the same points were used, except for eleven tie points that were collimated only in four images in those areas close to borders where not all the images overlapped. The overall root mean square error (RMSE) of the residuals was well below one pixel.

The entire photogrammetric process was performed in the OrthoEngine tool of the Geomatica suite (version 2018, PCI Geomatics, Markham, ON, Canada). The adopted matching strategy is based on the semi-global matching technique. The output is composed by a digital elevation model for each stereo-couple and a score channel, which classifies elevation values in three levels of confidence, depending on the goodness of the matching.

Considering the complex morphology of urban areas, an extremely high level of details is required to correctly describe the shape of buildings and infrastructures. For this reason, the DEM extraction was performed with the highest possible resolution, which is the original resolution of the panchromatic band (0.3 m), and no smoothing filters were applied. These choices allow for testing the potentials of these kinds of images. Furthermore, not only the three original stereo-couples were matched for the DEM generation, but also all the other possible combinations of the six images were considered;

therefore, a total of fifteen DEMs were generated and evaluated with the reference data (Table 2). This approach allows for testing the effects of different acquisition geometries. In fact, the visibility of roads and urban canyons varies significantly depending on the satellite position. For this purpose, the convergence angle, the bisector elevation angle (BIE) and the asymmetry angle were computed for every pair. Calculations are based on the azimuth and the elevation angles provided in the WV3 metadata, according to the formulas available in the literature [19,20]. These three angles are used to measure the geometrical relationship between the two optical rays (one from the fore image and one from the aft image) that intersect at a common ground point. Convergence, indeed, is the angle between the two rays in the epipolar plane and reflects the base-to-height ratio; BIE is defined as the elevation angle of the bisector of the convergence angle and describes the obliqueness of the epipolar plane; finally, asymmetry is measured between the bisector and the line perpendicular to the baseline and specifies the level of symmetry between the fore and aft observation rays [21]. The asymmetry angles, however, could be computed only with rough approximation using the nominal height of the satellite because the exterior orientation parameters are not delivered.

Table 2. List of the WorldView-3 pairs used for DEM extraction and angles describing the acquisition geometry. The first three are the original pairs from the acquisition.

Pair		Convergence	BIE	Asymmetry
2938	2824	46.6	71.4	10.2
2526	2626	37.5	66.1	8.8
2549	2458	31.6	65.2	10.5
2824	2458	8.6	54.9	10.3
2938	2526	24.4	71.0	3.4
2938	2626	15.4	62.3	26.3
2824	2526	23.8	61.3	16.5
2824	2549	38.7	66.1	10.8
2824	2626	61.0	70.4	0.9
2938	2458	40.6	70.0	9.5
2458	2526	16.5	61.5	18.6
2458	2626	54.1	67.5	0.5
2526	2549	15.1	67.6	2.0
2549	2626	22.5	62.5	16.4
2938	2549	10.8	68.8	5.1

Finally, an average DEM was generated (Figure 3), in which the elevation value in each pixel was computed as a weighted average of the fifteen values measured in the stereo-couples, using the score channels to derive weights. To refine the solution, the weighted standard deviation was also computed and the height values falling outside a defined confidence interval (1.5 times the standard deviation) were discarded.

3.2. Validation

For each of the extracted DEMs, the following analyses were carried out. Firstly, the completeness was evaluated as the percentage of pixels successfully matched by the algorithm. Secondly, the computed elevations were compared with NRTK data at the locations of the 30 control points. Finally, the models were compared with the point clouds obtained from LiDAR and MIDAS5 images, using the Multiscale Model to Model Comparison plug-in (M3C2), implemented in the open-source CloudCompare software package (version 2.10). Lague et al. [22] present a thorough explanation of this method, which calculates the local distance between two point clouds (one working as reference). For any given point of the reference cloud, the distance is measured either along the vertical or along the direction of the local surface normal, which is determined by fitting a plane to the neighbours of that cloud. According to Gomez Gutierrez et al. [23], the M3C2 algorithm permits measuring the distance between two point clouds in a more realistic way than other methods based on raster DEMs or 2.5D grids, providing the most accurate estimation of the point cloud quality.

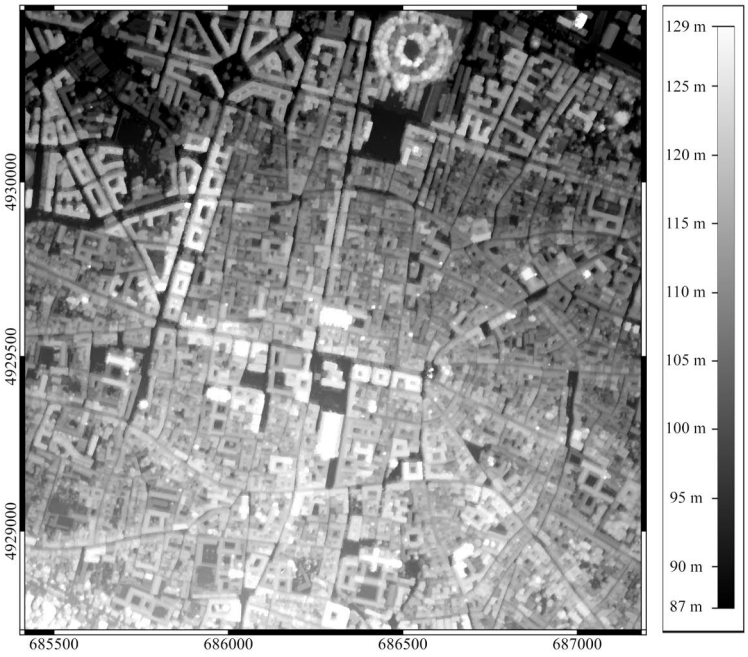


Figure 3. Grey-level visualization of the Average DEM of the city centre.

Referring to Figure 1, the comparison based on M3C2 was carried out in three smaller areas that present different urban textures. The first area (called Centre, dimensions 1.8 × 1.8 km) is located in the city centre, where building density is higher and roads are very narrow; the second one (called Corticella, dimensions 0.8 × 0.7 km) is a residential area located in the Corticella district, with wider avenues and more open spaces, but also with modern buildings with less regular shapes. The last one (named Fair, dimensions 1.7 × 1.6 km) was chosen in a productive district and includes some warehouses used for exhibitions and fairs. In Centre and Fair areas, LiDAR data were available for the comparisons, whilst, in Corticella zone, the MIDAS5 model was used.

For the computation of M3C2 distance, all of the models were converted in LAS format and subsampled, when necessary, to ensure an almost equal point density. The M3C2 distance for each point was measured along the vertical [24] and the diameter of the search cylinder was set to twice the average point density (2 m for comparisons with LiDAR and 0.7 m with MIDAS5). Overall statistics (including means, quartiles, 5th and 95th percentiles) were computed in the three areas and separate statistics for building roofs and road carriageways were also computed. The points belonging to these two urban features were selected in QGIS environment, by means of an intersection operation between the point clouds and the polygons of the technical cartography. A buffer of 0.3 m (corresponding to one WV3 pixel) inside the polygons of buildings and carriageways was considered, in order to prevent possible co-registration errors.

4. Results and Discussion

The results in terms of completeness of the whole model are summarised in Figure 4a. Considering the models derived by the individual stereo-pairs, the best solution is given by the 2526–2549 couple that reaches completeness of about 89%, while the less complete solution is given by the 2824–2626 couple (48%). An inverse correlation was found between completeness and convergence angle, with a coefficient of determination (R^2) equal to 0.8, if the entire 100 km^2 area is considered. A similar trend was observed also by Aguilar et al. [12], even though with a lower R^2, and seems to be typical of urban areas. The comparison of the three validation areas suggests a connection with building density: R^2, indeed, rises up to 0.90 in Fair area and even 0.97 in the Centre, whilst drops to 0.69 at Corticella. The success of the matching is highly dependent on the visibility of the mapped surfaces. Obviously, tall buildings hide many façades and portions of carriageways; this problem is exacerbated

by the lack of a really nadiral image (as for Table 1, off-nadir angles range from 18 to 33 degrees), which is very common for satellite imagery.

The Average DEM (which was obtained as a weighted average of the fifteen DEMs from the single stereo-pairs, Section 3.1) took advantage of the different view points of all the six images and reached completeness of 99%. Many tests were carried out to assess the achievable completeness with an increasing number of images, by merging the DEMs computed from all the possible pairwise combinations of only three, four and five images, respectively. Results in Figure 4b show a monotonic but asymptotic growth of the average completeness with the number of images available and a simultaneous narrowing of the variability (clearly, the range for the completeness with six images is not drawn because only one solution—which is the Average DEM of Figure 4a—is possible with the available dataset).

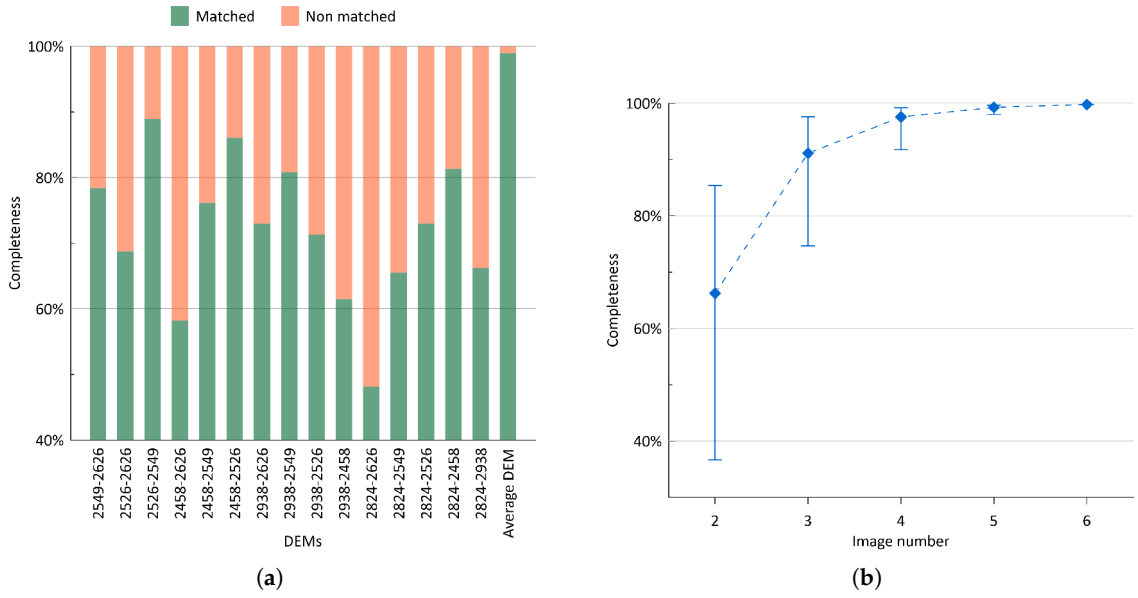


Figure 4. (**a**) completeness of the entire DEMs derived from each pair and of the Average DEM; (**b**) completeness variation with the number of images available for the Centre area.

Statistics on check points are shown in Figure 5. RMSEs on the elevations are in the range of two pixels for most of the extracted DEMs, except for four cases where few outliers rise the variance. These outliers are related to severe matching problems in correspondence of some ancient towers in the historical centre of the city. Towers are structures characterised by a very limited extent in planimetry, but extremely pronounced height. An example of this situation is given in Figure 6, where the computed elevations of the Garisenda tower are compared with reference data. Looking at the screenshots of the six WV3 images, it appears that the matching algorithms are not able to manage strong differences in the perspective view. For instance, when the pair is made of two images in which different façades of the tower are visible (e.g., 2824–2549), the matching fails completely and the elevation is not computed. On the other hand, when the perspective view of the tower is almost the same (e.g., 2938–2549), the elevation is correctly computed. The worst cases, however, are the intermediate ones (i.e., 2938–2526, 2938–2458, 2824–2938), where strongly biased values are obtained and the score does not highlight the anomaly. Figure 6 demonstrates also the importance of a rejection criterion in the computation of the Average DEM.

Statistics on M3C2 distances calculated in the three areas between the WV3 Average DEM and the reference data (LiDAR for Centre and Fair, and MIDAS5 for Corticella) are shown in Table 3. The values are substantially comparable in the three areas, except for the median in Corticella, which is higher, probably because of the more complex shape of the roofs (which are composed of parts with different heights). On the other hand, in this area, the range of M3C2 values is narrower (from −30 to 37 m,

instead of −74 to 71 m in the city centre) because there are not particularly problematic structures, such as towers.

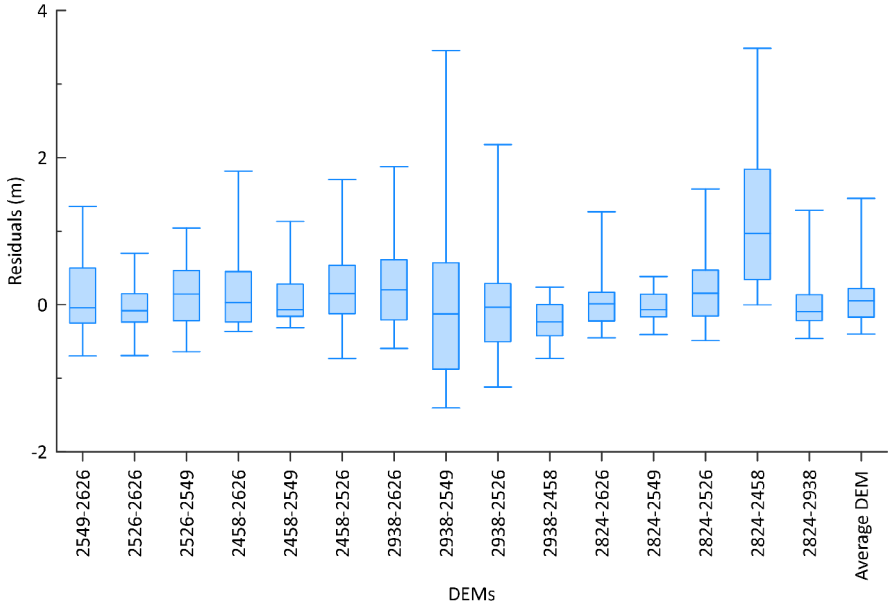
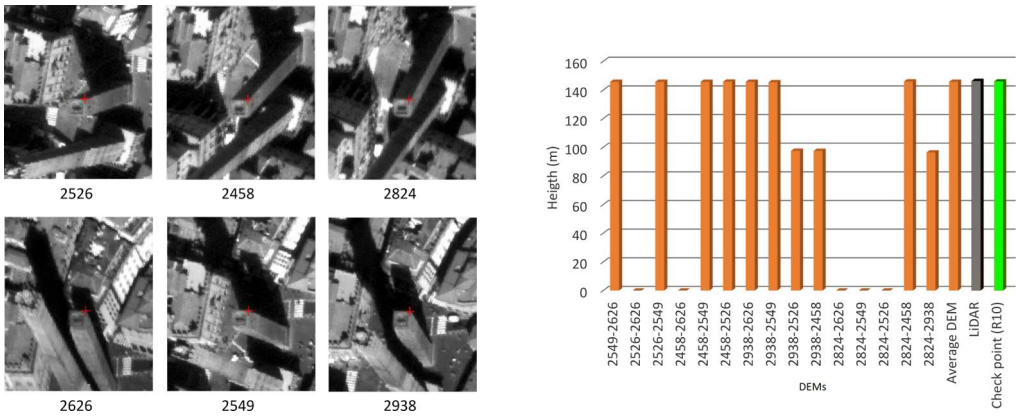


Figure 5. Statistics on the check point residuals for the DEMs derived from each pair and for the Average DEM. Box limits are the first and third quartiles, while the band inside is the median. Whiskers represent the 5th and 95th percentiles.



(**a**) (**b**)

Figure 6. Reconstruction of the Garisenda tower. (**a**) screenshots of the six WV3 images of the tower; (**b**) computed heights of the tower according to the WV3 DEMs and reference data.

Table 3. Statistics on the M3C2 distance (in metres) computed between the WV3 Average DEM and the reference data in the three areas.

	Centre (WV3-Lidar)			Fair (WV3-Lidar)			Corticella (WV3-MIDAS5)		
	Overall	**Roofs**	**Roads**	**Overall**	**Roofs**	**Roads**	**Overall**	**Roofs**	**Roads**
Points	3,258,420	1,598,139	617,660	2,711,801	555,437	427,552	14,637,983	1,258,766	2,020,506
Mean	2.76	0.61	6.06	1.83	0.93	1.55	3.40	3.27	3.58
STD	5.39	3.19	6.65	4.91	4.54	4.13	5.18	5.29	5.08
Min	−74.04	−74.04	−38.47	−81.86	−58.46	−65.49	−30.35	−30.35	−22.98
5th perc.	−2.22	−1.73	−2.62	−2.82	−1.66	−2.15	−2.12	−0.5	−1.56
Median	0.34	−0.22	5.38	0.25	-0.09	0.18	1.72	1.10	1.70
IQR	5.65	0.98	10.21	3.11	0.68	2.26	5.07	3.46	4.97
95th perc.	13.84	7.05	17.61	10.89	9.17	10.21	13.69	14.53	13.90
Max	71.93	52.78	71.93	79.00	69.24	35.67	37.00	31.76	37.00

The M3C2 map between the Average DEM and LiDAR for the Centre area is shown in Figure 7. It is evident that the more problematic parts are the streets, where the elevation is always overestimated. This is also the reason for the marked asymmetry of the distributions toward positive values. In order to quantify this phenomenon, separate calculations of M3C2 statistics were computed for building roofs and streets. Results confirm a far better accuracy for roofs (Figure 8 and Supplementary Materials), with medians between −0.3 and 0.1 m and interquartile ranges (IQRs) between 0.8 and 1.1 m, compared to roads (medians 2.5–6.4, IQRs 8.4–11.7 m), even though some outliers still influence mean values (maximum M3C2 values for roofs are in the range 33.9–70.9 m, whilst the 95th percentiles are 5.5–8.1 m). A section of the point clouds in correspondence of a street in the city centre is shown in Figure 9b. The analysis of the profiles makes evident the inability of the generated DEM in describing correctly street shapes and, more generally, vertical façades, resulting in the local overestimation of the ground level. This is particularly apparent in the narrow urban canyon of the city centre and is confirmed also for Fair area, even if less evident, probably because of wider roads and more visible carriageways. In Corticella, this phenomenon is instead not so clear (statistical parameters are similar for roofs and roads) because, on the one hand, buildings are more spaced and the number of exposed façades is higher; on the other hand, frequent height differences between roof parts pose the same reconstruction problems of façades. Here, the biggest differences occur in a buffer zone around buildings because the extracted DEMs are not able to reconstruct the shape of vertical walls (which are instead well defined in the reference model) and tend to connect ground and roof levels with slopes (Figure 9d). When limiting the analysis to the central portion of the roofs with simpler shapes, the statistics on M3C2 distance tend to conform to the values of the other zones.

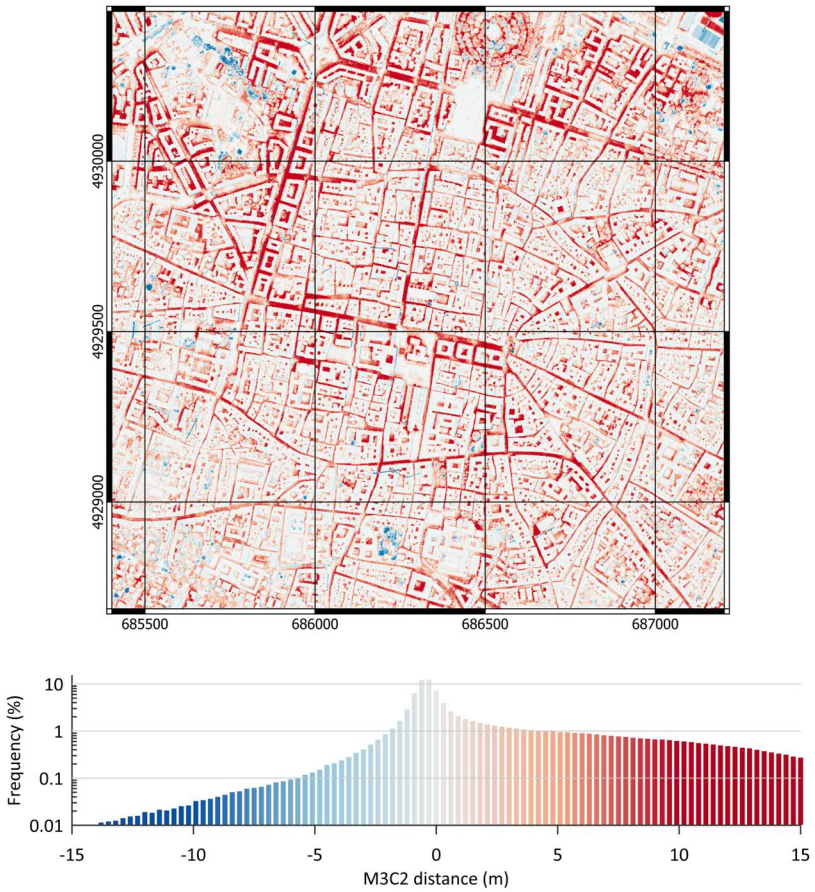
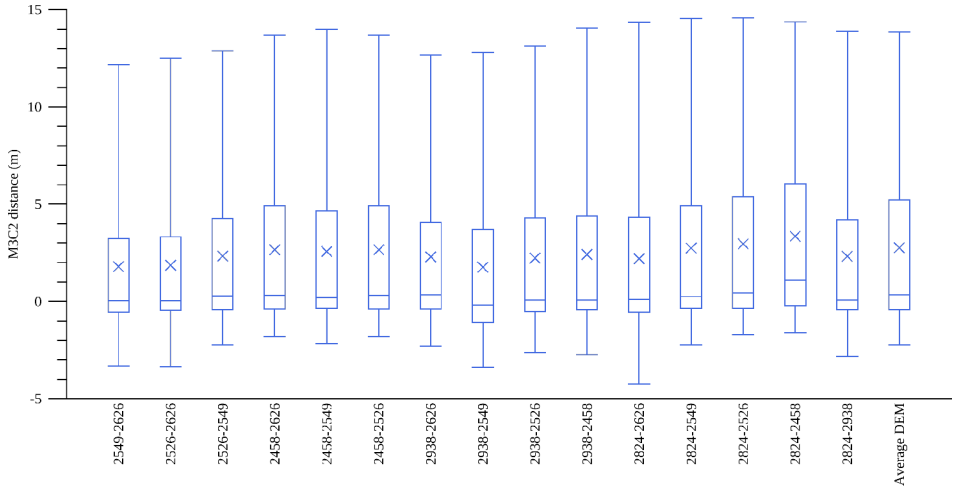
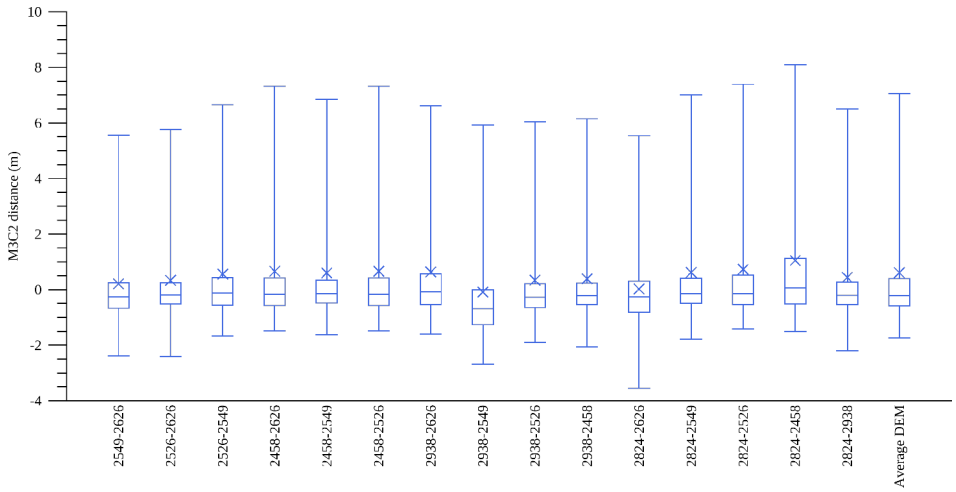


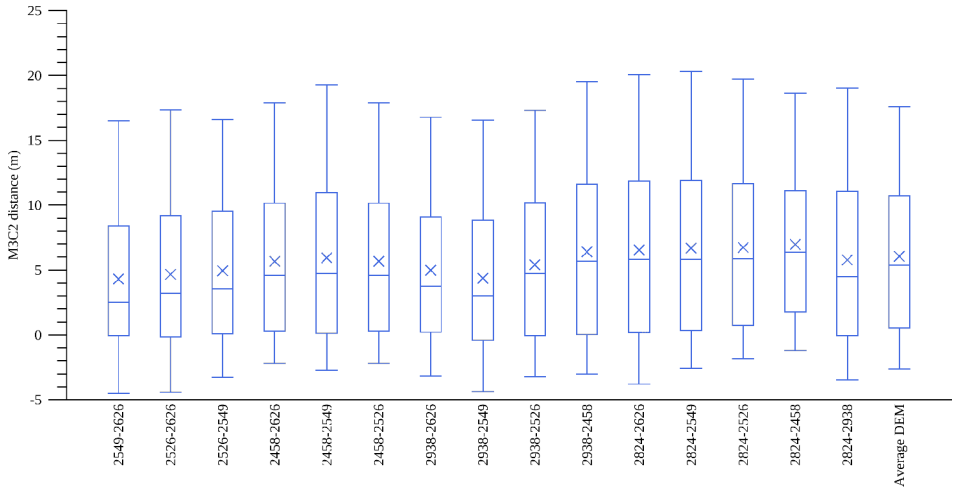
Figure 7. Map and histogram of the M3C2 distance between the Average DEM and the reference LiDAR data in the Centre area.



(**a**)



(**b**)



(**c**)

Figure 8. Statistics on M3C2 distance in the Centre area for all the computed DEMs, using LiDAR data as reference. Box limits are the first and third quartiles, while the band inside is the median. Whiskers represent the 5th and 95th percentiles. Crosses represent means. (**a**) statistics on the whole area; (**b**) on roof surfaces; (**c**) on carriageway surfaces. Values are reported in the Supplementary Materials.

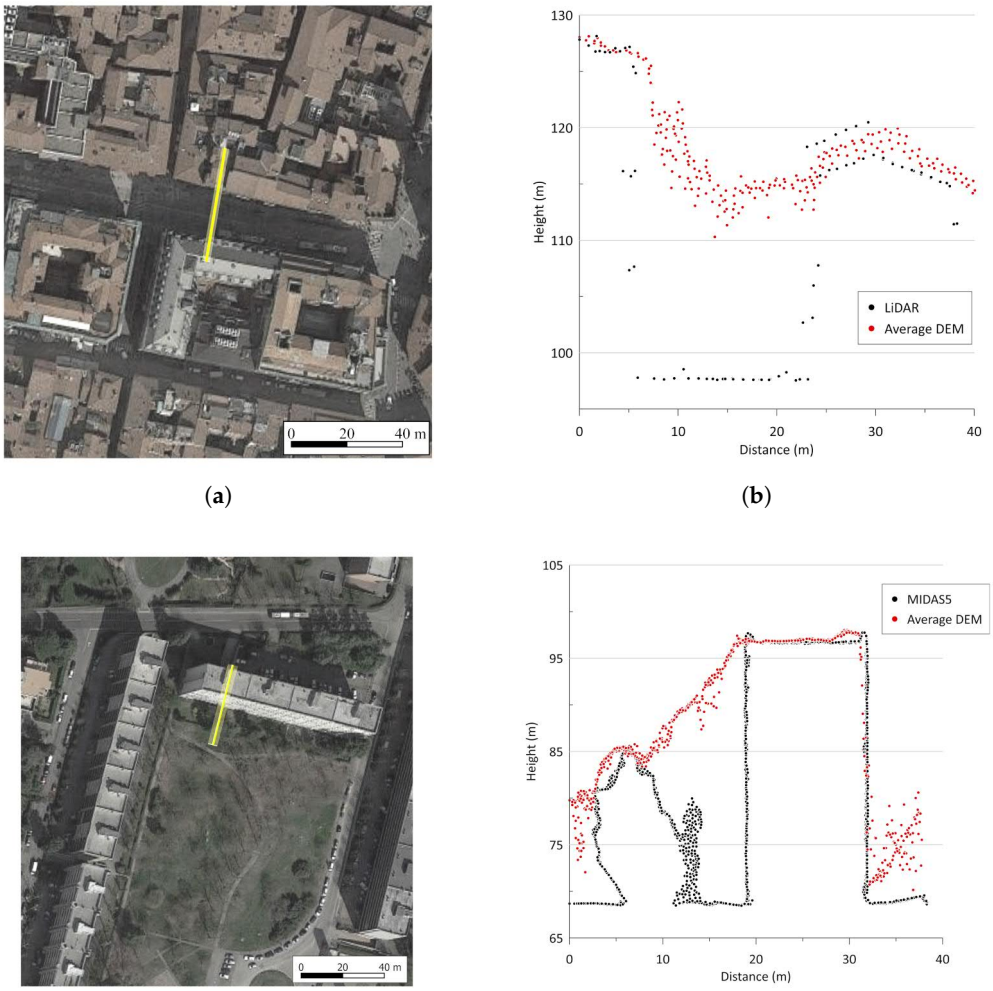


Figure 9. Sections of the point clouds extracted in the Centre area (**a**,**b**) and in Corticella area (**c**,**d**), respectively. A buffer of 1.5 m was considered to select points.

The same trends are confirmed in the DEMs generated by the single stereo-pairs, where IQR ranges from 3.8 m to 6.2 m and are again far better for roof surfaces only (Figure 8). It is to be noticed that some stereo-pairs show slightly better statistics compared to the Average DEM. These particular pairs seem to show an optimal compromise between two opposite needs: framing the objects so that the two images present a good similarity level, which is important for the image-matching algorithm, and a not too short baseline, which influences, as known, the standard deviation of the Z-coordinate for the extracted points. For example, the 2824–2458 pair has a convergence angle of 8.6 degrees and a median on the M3C2 of 1.09 m, while the 2938–2549 pair has a convergence of 10.8 and a median of −0.19. Nevertheless, the IQRs are similar and dominated by errors on roads, which are always present to a certain extent because, as can be seen in Figure 2, none of the images is well aligned with the most frequent directions of roads in the study area. Overall, significant correlations between medians, IQRs and the angles describing the acquisition scheme were not found (R^2 coefficients are reported in Table 4). This is in contrast with the findings of other authors who reconstructed areas with different land covers [25], probably because, in urban environments, statistics are dominated by local matching errors.

Table 4. Coefficients of determination (R^2) between M3C2 median, M3C2 IQR, DEM completeness and the angles defining the acquisition geometry. Correlations are based on the values obtained from the fifteen models of the Centre zone.

	Overall			Roofs			Roads		
	Median	**IQR**	**Completeness**	**Median**	**IQR**	**Completeness**	**Median**	**IQR**	**Completeness**
Convergence	0.10	0.02	0.97	0.00	0.21	0.98	0.08	0.41	0.92
BIE	0.33	0.18	0.29	0.28	0.36	0.30	0.02	0.21	0.24
Asymmetry	0.05	0.00	0.14	0.09	0.00	0.19	0.01	0.07	0.06

The availability of more than one stereo-couple appears important in terms of completeness (which is one of the major concerns in urban areas) and in order to detect outliers in the final products (which remain often undetected in the analysis of the matching score provided by the adopted photogrammetric software).

Considering the relationships between accuracy and the satellite azimuth of the acquisitions, it can be noted that the best results (median and IQR of the M3C2 distance, Figure 10) are obtained with pairs including image 2549 (2549–2626, 2938–2549, 2526–2549), whose azimuth is 115 degrees and is the closest to the most frequent road directions (Figure 2b). In addition, 2526–2626 presents similar results for M3C2 distance, but the completeness is lower (68% versus 78–89%). Nevertheless, a clear trend cannot be inferred in this case study, probably because a direct comparison of the pairs only in terms of azimuth is not fully explanatory due to the simultaneous variation of all the other acquisition geometry parameters.

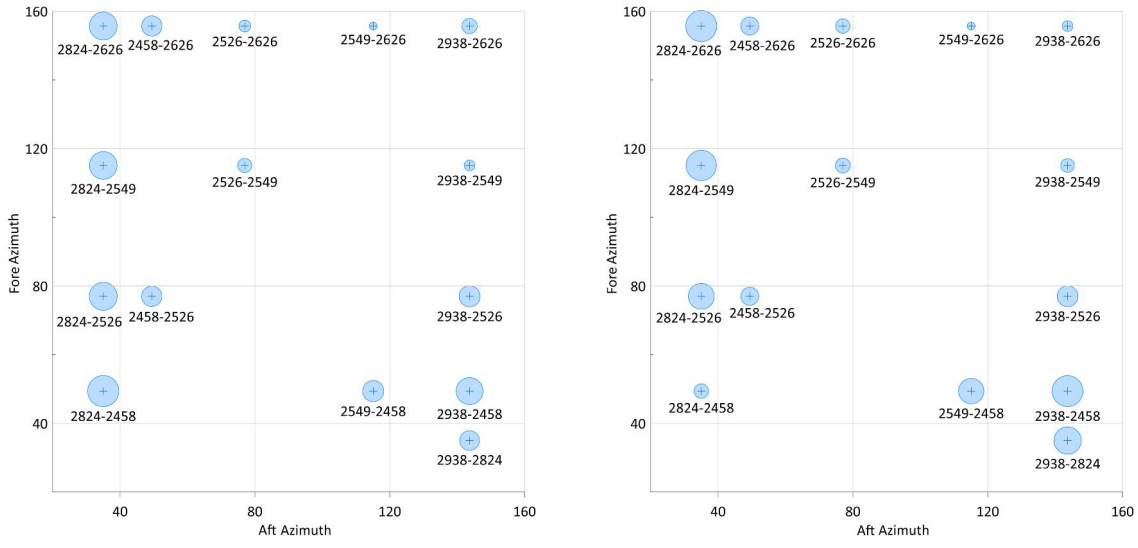


Figure 10. Medians (**a**) and IQRs (**b**) of the M3C2 distance in the Centre area depending on the azimuths of the fore and aft images of stereo-pairs. The size of the bubbles is proportional to median and IQR value, respectively.

All in all, the results of the experimentations suggest that elevation models from WV3 are sufficiently reliable for applications where the overall elevation of urban features is more important than the reconstruction of the exact shape (such as orthorectifying high-resolution multispectral images or enriching the information content of municipal technical cartographies about buildings [26]). On the other hand, WV3 models do not meet the quality requirements for more detailed 3D modelling of cities (targeted, for example, at the compilation of 3D cadastres or the development of emergency response plans [27]).

5. Conclusions

The experiments carried out highlighted potentials and limitations of the automatic generation, in an ordinary productive scenario, of digital elevation models in urban areas from WV3 images, which offer nowadays the highest spatial resolution among the commercially available sensors with stereo capabilities.

The weak acquisition geometry is the main limiting factor for an accurate description of the urban texture. With a single stereo-pair, the achievable completeness varies between 50% and 90%. Better results can be obtained using more than one couple (using three images the average completeness doubles, with six images reaches 99%), even though the costs of this imagery can be a constraint. A clear negative correlation (R^2 up to 97%) with the convergence angle is observed where building density is high, thus pairs with smaller convergence angles (in the range 8–16 degrees) perform better. This apparent anomaly can be explained considering the nature of the matching process, which can be completely hampered by excessive changes in the perspective view of buildings. For this reason, the acquisition of two single images can sometimes provide better results, compared to the acquisition of a standard stereo-pair product, because this product usually provides higher convergence angles (32–47 degrees in the presented dataset).

The accuracy analyses, separately performed on different components of the urban textures, highlighted an average difference between WV3 models and reference ones of about 0.6 m with a standard deviation of about 3 m on roof surfaces. However, strong biases were observed on roads and adjacent to building façades (mean difference 6.0 m, standard deviation 6.7 m).

As a final consideration, for applications where the overall elevation of urban features is more important than the precise reconstruction of the shape, the quality of the DEMs from WorldView-3 images can be considered adequate. Nevertheless, in urban areas, the accuracy is still not competitive with the one achievable with other techniques, such as LiDAR, especially when considering the capability of reconstructing vertical surfaces. Better results will perhaps be achievable when experimental algorithms proposed in the scientific literature will be implemented in the production workflow.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/11/7/878/s1, Table S1: Full statistics on M3C2 distance in the Centre area.

Author Contributions: E.M.: conceptualisation, funding acquisition and DEM extraction. V.A.G.: M3C2 computations, validation and paper editing. L.P.: GNSS surveys and validation.

Funding: This work is part of the ChoT project, funded by the Italian Ministry of Education, University and Research in the framework of the Scientific Independence of young Researches (SIR) 2014 programme (grant number RBSI14SYES).

Acknowledgments: The authors would like to thank Andrea Minghetti, Paola Africani and Elisa Paselli of the "SIT" division of the Municipality of Bologna, for providing cartography, LiDAR data and MIDAS5 imagery.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

BIE	Bisector Elevation Angle
DEM	Digital Elevation Model
GSD	Ground Sample Distance
IQR	Inter-Quartile Range
MIDAS5	Multi-camera Integrated Digital Acquisition System
M3C2	Multiscale Model to Model Comparison
NRTK	Network Real-Time Kinematic
RMSE	Root Mean Square Error
WV3	WorldView-3

References

1. Longbotham, N.; Pacifici, F.; Baugh, B.; Camps Valls, G. Prelaunch assessment of WorldView-3 information content. In Proceedings of the 2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lausanne, Switzerland, 24–27 June 2014. [CrossRef]
2. Satellite Imaging Corporation. WorldView-3 Satellite Sensor. Available online: https://www.satimagingcorp.com/satellite-sensors/worldview-3/ (accessed on 31 January 2019).
3. Barazzetti, L.; Roncoroni, F.; Brumana, R.; Previtali, M. Georeferencing accuracy analysis of a single WorldView-3 image collected over Milan. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B1*, 429–434. [CrossRef]
4. Sidike, P.; Sagan, V.; Maimaitijiang, M.; Maimaitiyiming, M.; Shakoor, N.; Burken, J.; Mockler, T.; Fritschi, F.B. dPEN: Deep progressively expanded network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery. *Remote Sens. Environ.* **2019**, *221*, 756–772. [CrossRef]
5. Mars, J.C. Mineral and lithologic mapping capability of WorldView 3 data at Mountain Pass, California, using true- and false-color composite images, band ratios, and logical operator algorithms. *Econ. Geol.* **2018**, *113*, 1587–1601. [CrossRef]
6. Zhang, P.; Ke, Y.; Zhang, Z.; Wang, M.; Li, P.; Zhang, S. Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery. *Sensors* **2018**, *18*, 3717. [CrossRef] [PubMed]
7. Parente, C.; Pepe, M. Bathymetry from WorldView-3 satellite data using radiometric band ratio. *Acta Polytech.* **2018**, *58*, 109. [CrossRef]
8. Aguilar, M.A.; Nemmaoui, A.; Aguilar, F.J.; Qin, R. Quality assessment of digital surface models extracted from WorldView-2 and WorldView-3 stereo pairs over different land covers. *GISci. Remote Sens.* **2018**, *56*, 109–129. [CrossRef]
9. Hu, F.; Gao, X.M.; Li, G.Y.; Li, M. DEM extraction from WorldView-3 stereo-images and accuracy evaluation. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B1*, 327–332. [CrossRef]
10. Capaldo, P.; Crespi, M.; Fratarcangeli, F.; Nascetti, A.; Pieralice, F. DSM generation from high resolution imagery: Applications with WorldView-1 and Geoeye-1. *Ital. J. Remote Sens.* **2012**, *44*, 41–53. [CrossRef]
11. Rupnik, E.; Pierrot Deseilligny, M.; Delorme, A. 3D reconstruction from multi-view VHR-satellite images in MicMac. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 201–211. [CrossRef]
12. Aguilar, M.A.; del Mar Saldana, M.; Aguilar, F.J. Generation and quality assessment of stereo-extracted DSM from GeoEye-1 and WorldView-2 imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1259–1271. [CrossRef]
13. Hirschmuller, H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 328–341. [CrossRef] [PubMed]
14. Bethmann, F.; Luhmann, T. Semi-global matching in object space. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *XL-3/W2*, 23–30. [CrossRef]
15. Topcon Positioning Italy. NetGEO. Available online: http://www.netgeo.it/ (accessed on 31 January 2018).
16. Dabove, P.; De, M.; Manzino, A. Achievable positioning accuracies in a network of GNSS reference stations. In *Global Navigation Satellite Systems: Signal, Theory and Applications*; InTech: London, UK, 2012. [CrossRef]
17. Africani, P.; Bitelli, G.; Lambertini, A.; Minghetti, A.; Paselli, E. Integration of LiDAR data into a municipal GIS to study solar radiation. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *XL-1/W1*, 1–6. [CrossRef]
18. Bitelli, G.; Girelli, V.A.; Lambertini, A. Integrated use of remote sensed data and numerical cartography for the generation of 3D city models. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII-2*, 97–102. [CrossRef]
19. Li, R.; Niu, X.; Liu, C.; Wu, B.; Deshpande, S. Impact of imaging geometry on 3D geopositioning accuracy of stereo Ikonos imagery. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 1119–1125. [CrossRef]
20. Jeong, J.; Kim, T. Quantitative estimation and validation of the effects of the convergence, bisector elevation, and asymmetry angles on the positioning accuracies of satellite stereo pairs. *Photogramm. Eng. Remote Sens.* **2016**, *82*, 625–633. [CrossRef]
21. Jeong, J.; Kim, T. Analysis of dual-sensor stereo geometry and its positioning accuracy. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 653–661. [CrossRef]

22. Lague, D.; Brodu, N.; Leroux, J. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 10–26. [CrossRef]
23. Gómez Gutiérrez, A.; de Sanjosé Blasco, J.; Lozano Parra, J.; Berenguer Sempere, F.; de Matías Bejarano, J. Does HDR pre-processing improve the accuracy of 3D models obtained by means of two conventional SfM-MVS software packages? the case of the Corral del Veleta rock glacier. *Remote Sens.* **2015**, *7*, 10269–10294. [CrossRef]
24. Conte, P.; Girelli, V.A.; Mandanici, E. Structure from motion for aerial thermal imagery at city scale: Pre-processing, camera calibration, accuracy assessment. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 320–333. [CrossRef]
25. Piermattei, L.; Marty, M.; Karel, W.; Ressl, C.; Hollaus, M.; Ginzler, C.; Pfeifer, N. Impact of the acquisition geometry of very high-resolution Pléiades imagery on the accuracy of canopy height models over forested alpine regions. *Remote Sens.* **2018**, *10*, 1542. [CrossRef]
26. Franci, F.; Lambertini, A.; Bitelli, G. Integration of different geospatial data in urban areas: A case of study. In *Proceedings Volume 9229, Second International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2014), Paphos, Cyprus, 12 August 2014*; Hadjimitsis, D.G., Themistocleous, K., Michaelides, S., Papadavid, G., Eds.; SPIE: Bellingham, WA, USA, 2014. [CrossRef]
27. Biljecki, F.; Stoter, J.; Ledoux, H.; Zlatanova, S.; Çöltekin, A. Applications of 3D City Models: State of the Art Review. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 2842–2889. [CrossRef]