Quality analysis of worldview-4 DSMS generated by least squares matching and semi-global matching

Umut G. Sefercik,[a] Mehmet Alkan,[b,*] Karsten Jacobsen,[c] Can Atalay,[d] Gurcan Buyuksalih,[e]

[1] Gebze Technical University, Department of Geomatics Engineering, 41400, Kocaeli, Turkey, sefercik@gtu.edu.tr
[2]Yildiz Technical University, Department of Geomatics Engineering, 34220, Istanbul, Turkey, alkan@yildiz.edu.tr
[3]Leibniz University Hannover, Institute of Photogrammetry and Geoinformation, Nienburger Str. 1, 30167, Hannover, Germany, jacobsen@ipi.uni-hannover.de
[4]Zonguldak Bulent Ecevit University, Department of Geomatics Engineering, 67100, Zonguldak, Turkey, atalay@beun.edu.tr
[5]Private Sector, Istanbul Metropolitan Municipality, 34430, Beyoglu, Istanbul, Turkey, gbuyuksalih@yahoo.com

Abstract. For the estimation of a dense disparity map from a rectified space-borne stereo image pair and generation of the most qualified digital surface model (DSM), new image matching algorithms are being developed permanently. This study has two main goals as the DSM quality validation of Worldview-4 (WV-4), which offer the highest ground sampling distance (GSD) (30 cm) of civilian optical space-borne missions together with Worldview-3 (WV-3) and performance comparison of Semi-Global Matching (SGM) and Least Squares Matching (LSM), two of the most preferred image matching algorithm for space-borne data. In the Istanbul study area with a rough topographic structure, 1 m gridded DSMs were derived from geometrically corrected WV-4 stereo image pair. The qualities were estimated by well-rounded visual and statistical approaches based on standard deviation and normalized median absolute deviation. In model-to-model comparisons, a very high-resolution airborne laser scanning (ALS) DSM was utilized as the reference model. The results demonstrated that the WV-4 DSM derived by SGM has the standard deviation of ±0.51 m (1.7 GSD) in height in open areas, where the LSM DSM has ±1.46 m. Besides, the visual quality of SGM is much better than LSM through building description potential.

Keywords: Worldview-4, least squares matching, semi-global matching, digital surface model, Quality.

***Second Author, E-mail: alkan@yildiz.edu.tr

1 Introduction

In recent years, by force of rapid developments in remote sensing technology, high-resolution digital surface models (DSMs), digital three dimensional (3D) cartographic representation of all-natural and human-made objects in the target area, became the most beneficial and indispensable products in a large variety of applications such as 3D city modelling[1], the establishment of geographical information systems (GIS)[2], fusion approaches in space-borne automatic object extraction[3], disaster monitoring and management[4], generation of forest stand height maps[5] and detection and monitoring of archaeological and cultural heritage sites.[6]

Spatial resolution is the most significant parameter to describe the 3D visible surface of the DSM's target area as close as to the real surface. In the last decade, with valuable contributions of DigitalGlobe, the spatial resolution of optical space-borne missions was improved and "very high resolution (VHR ≤ 0.5 m)" concept entered space-borne remote sensing technical glossary in 2007 with Worldview-1 mission.[7] Beginning with WV-1, DigitalGlobe continued with the series of commercial Earth-imaging satellites increasing the spatial resolution and Worldview-4 (WV-4), the latest member of the Worldview series, was successfully launched in November 2016. WV-4 presents 0.31 m panchromatic (PAN) and 1.24 m multispectral (MS) imaging resolutions from 617 km orbital height at nadir view. In principle, VHR provides a qualitative leap in the space-borne missions' performance indexes such as geolocation accuracy and mapping capability.[8,9] The geolocation accuracy of WV-4 is specified as <5 m considering CE90 without ground control points (GCP) by DigitalGlobe. For high-quality mapping, the satellite has PAN+4 MS bands as red, green, blue and near-infrared (NIR) with an 11-bits dynamic range. To improve the geolocation accuracy and mapping quality of VHR satellite imagery, the literature mainly investigates the effect of image matching algorithms and different image orientation models such as deterministic models and rational function models (RFMs).[10-12] Accordingly, some researches exist about the ortho-rectification and DSM qualities of Worldview-2 (WV-2) and Worldview-3 (WV-3).[9,13-17] However, due to being the latest VHR Worldview mission, there is no comprehensive investigation about WV-4 geolocation accuracy and DSM quality in the literature. In this study, we had two aims: the comprehensive investigation of the WV-4 DSM potential and the comparison of least squares matching (LSM)[18] and semi-global matching (SGM)[19] effects on WV-4 DSM quality. The building monitoring performances of LSM and SGM DSMs generated from 1 m resolution IKONOS stereo imagery was compared in [20], and the

authors achieved better results with SGM. In [21], the effect of image-matching methods in DSM generation was also investigated with IKONOS data.

The paper is organized as follows: characteristics of the study area and analyzed WV-4 images are given in the next section. Image orientation, LSM and SGM techniques, DSM generation and validation methodologies are described in section 3. The achieved results are presented in section 4 and discussed in section 5, followed by the conclusions.

2 Characteristics of the Study Area and Analyzed WV-4 Images

Kilyos, a study area located in the northern part of Istanbul, has a shoreline with the Black Sea. Three classes dominate the land cover as built-up, open and forest. The land surface is rolling, and the orthometric heights are between 0 and 100 m. Fig. 1 shows the WV-4 image of the study area and land cover classes.

The characteristics and the acquisition geometry of used WV-4 optical stereo images are given in Table 1. The images have one minute eight seconds acquisition interval. WV-4 satellite has not captured new images since January 2019 because of a failure in its control moment gyros. However, the satellite's archive images are large in demand through the highest spatial resolution ever.

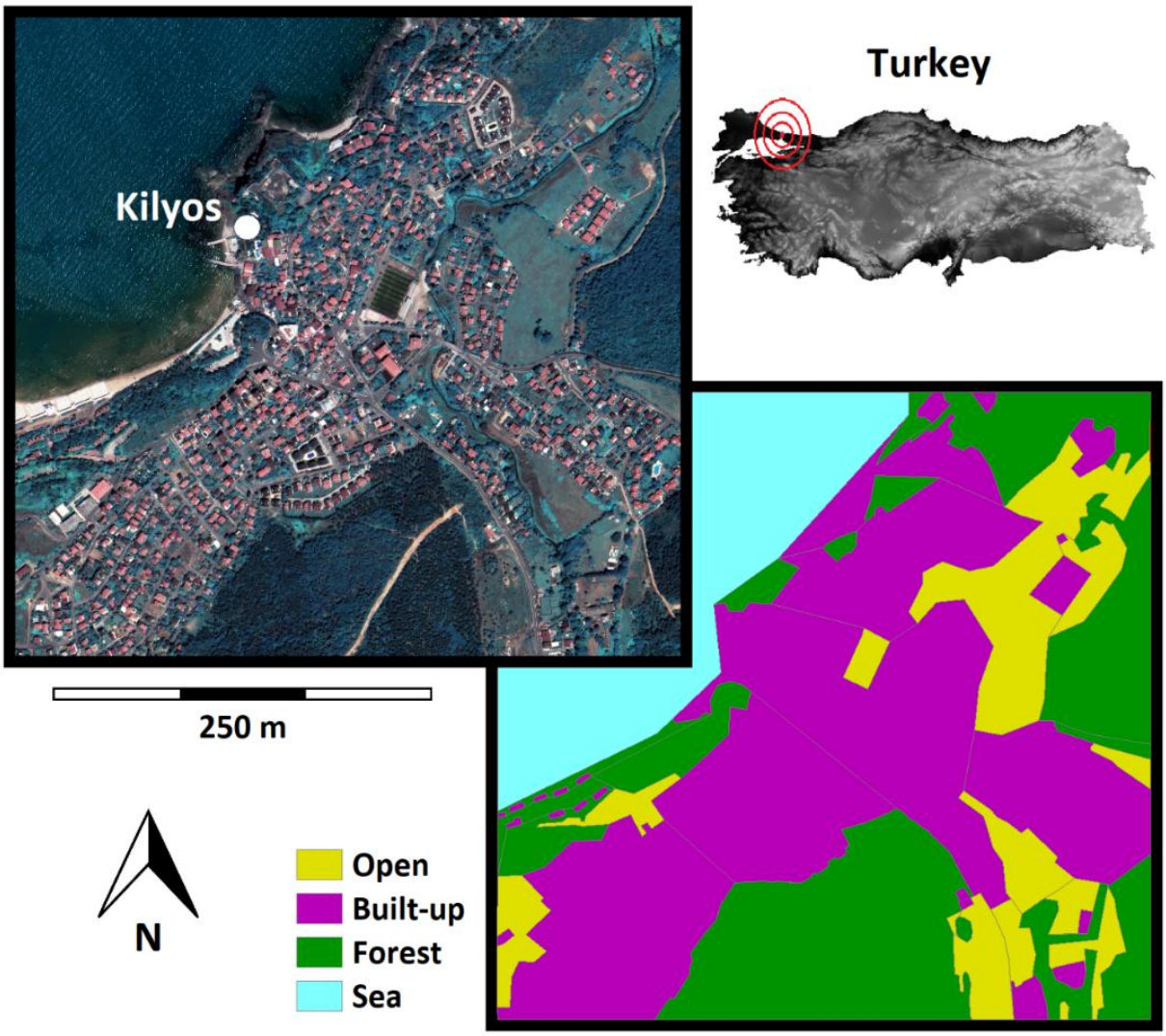
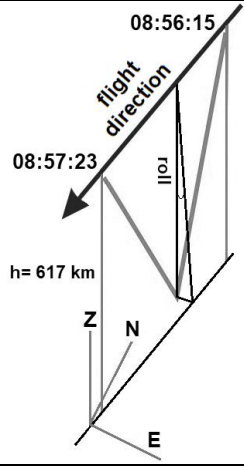


Fig. 1 WV-4 image of the study area and land cover classes

Table 1 Characteristics and the acquisition geometry of WV-4 stereo-images

Characteristics	Image 1	Image 2	Imaging Geometry
Used imaging band	PAN + MS + NIR		
Product type and mode	Stereo / Full swath		
Acquisition date	24.06.2017		
Acquisition time (h:m:s)	08:56:15.02	08:57:23.28	
Map projection and zone	UTM Zone 35		
Coverage	100 km^2		
Pixel size (Row×Column)	68616×25604 PAN / 17154×6401 MS		
Spatial resolution (at nadir)	0.31 m PAN / 1.24 m MS		
Radiometric level	Corrected		
Radiometric resolution	16 bit		
Scan direction	Forward		
Base length	375898 m		
Height to base relation	1 : 1.20		

3 Methodology

For the generation and validation of WV-4 DSMs, the following methodology steps, shown in Figure 2, were applied. Two WV-4 DSMs were generated with LSM and SGM algorithms and comprehensively analyzed separately. These processing steps are given in three sub-sections as image orientation, DSM generation by LSM and SGM and DSM validation.

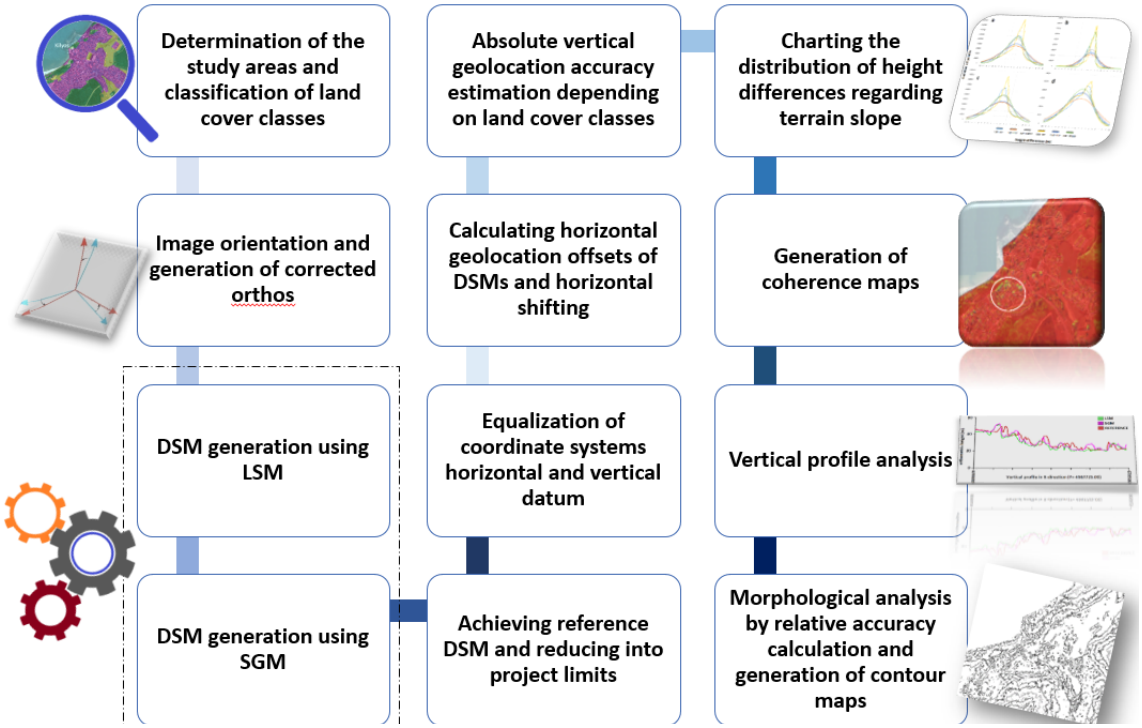


Fig. 2 Methodology steps.

3.1 Image Orientation

The DigitalGlobe specifies the horizontal and vertical geolocation accuracies of the Worldview series with circular error (CE90) and linear error (LE90) sequentially. “90” means that the minimum 90% of points measured have the geolocation error less than the stated value. The WV-4 images used in the study are stereo ortho-ready level 2A (OR2A), and both CE90 and LE90 were specified as < 5 m by DigitalGlobe for less than 30° off-nadir view and excluding terrain slope. To generate high accuracy large scale mapping products from WV-4 imagery,

inherently, the geometry needs an improvement that is usual for high resolution space-borne optical imagery.[22-24]

In this study, the bias-corrected sensor-oriented rational polynomial coefficient (RPC) model was used as the dominant method for high-resolution image orientation. RPC model enables direct mapping from 3D object space coordinates (usually offset normalized latitude, longitude, and height) to 2D image coordinates (usually offset normalized line and sample values) by incorporating image shift and drift terms.[25-28] RPCs express the relationship between the image and the ground coordinates and are based on the satellite's direct sensor orientation using positioning systems, such as global positioning systems, gyros, and star sensors, which determine the attitude data.[29,30] Together with the images and the header data, the RPCs were distributed, and seven (7) ground control points (GCPs) were established in the study area determined from 1⁄1000 scaled topographic vector maps generated from digital photogrammetry by Istanbul Metropolitan Municipality. In three-dimensional image orientation, the resulting standard deviation is 0.5 pixel in X, a 0.25 pixel in Y and 0.9 pixels in Z. The accuracy of image orientation is shown in Table 2, and the error vector plots are presented in Figure 3. In vector plots, the directions and the amounts of errors are not similar, demonstrating that the rectified ortho has no 3D systematic geolocation offset.

Table 1 The standard deviation of image orientation

RPC orientation of single images		
Image	SX	SY
Image 1	0.17 m	0.09 m
Image 2	0.22 m	0.25 m
Three-dimensional orientation		
Stereo model	SX	SY	SZ
	0.15 m	0.07 m	0.27 m

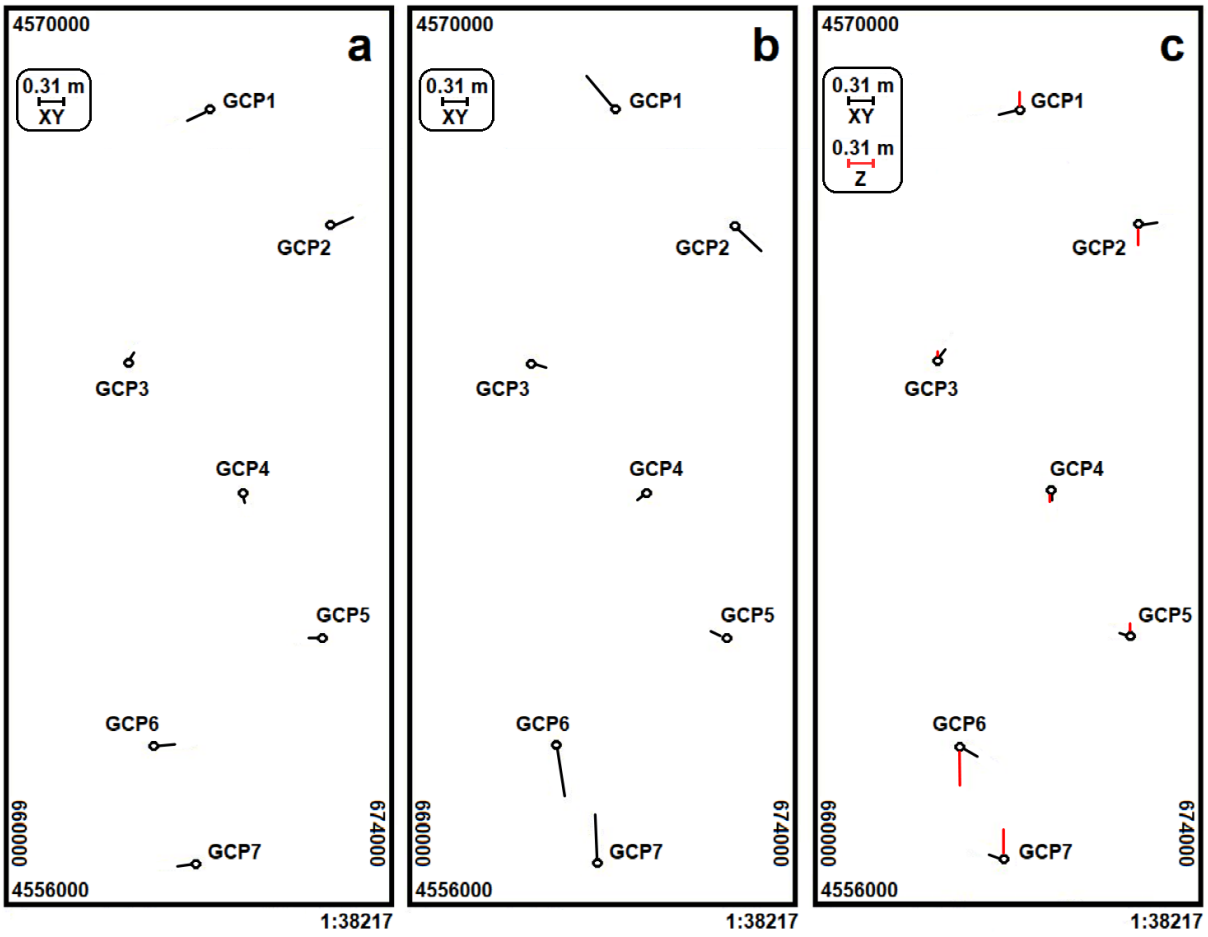


Fig. 3 The vector plots of the WV4 stereo imagery (a and b) and stereo model (c) based on bias-corrected RPC orientation [m] (red is the differences in Z, and upward/downward directions mean +/-).

3.2 DSM Generation by Least Squares Matching and Semi-Global Matching

In this study, one of the main objectives is to prove the effects of area-based LSM and pixel-based SGM algorithms on WV-4 DSM qualities. For high-quality DSM generation, accurate matching of the stereo images is crucial and depends on accurate image orientation. The image orientation accuracy should be less than 1 pixel; a lower accuracy leads to an increased number of wrong matches.[31] As shown in Table 2, orientation accuracy for the stereo model is 0.5 pixel in X, 0.25 pixel in Y and 0.9 pixels in Z, which indicate a suitable condition for matching.

LSM is one of the most preferred area-based image matching techniques in remote sensing since the 1980s. The basic principle of LSM is the geometric and radiometric matching of two or more image patches from a remotely-sensed master image concerning a slave image.[32] In 1982, Foerstner developed the LSM idea for a one-dimensional image line by integrating one translation parameter.[33] Then, the idea was adopted to the two-dimensional case by additional geometric and radiometric parameters using square or rectangular patches.[34-36] The method of LSM assumes that the best fit curve of a given type is the curve that has the minimal sum of deviations which is the least squares error from a given set of data. It can be defined as creating a curve that best fits several data points.

On the other hand, the process of creating mathematical functions is known as curve fitting. For a given dataset, assembly curves of a particular type are generally not unique. The optimum curve can be obtained with the least-squares method, with minimum deviation from all data points. If the dataset points are assumed as (x_1, y_1), (x_2, y_2),…, (x_n, y_n) and x is an independent variable and y is a dependent variable, the connection curve $f(x)$ has the deviation (error) e_i from each data point as given in equation 1.

$$\begin{aligned} e_1 &= y_1 - f(x_1) \\ e_2 &= y_2 - f(x_2) \\ &\vdots \\ e_n &= y_n - f(x_n) \end{aligned} \tag{1}$$

According to the LSM method, the equation of the best fit curve is given in equation 2. In this regard, the characteristic of the curve is minimal.

$$\sum_{1}^{n} e_i^2 = \sum_{1}^{n} [y_i - f(x_i)]^2 \tag{2}$$

SGM algorithm successfully combines the concepts of global and local stereo methods for accurate and pixel-wise matching at low runtime and estimates a dense disparity map from an oriented stereo image pair.[19,31,37] The key idea of SGM is to perform one-dimensional optimisations from all directions through the image regardless of sensor geometry. In the basic principle of SGM, two main components exist as matching cost computation and cost aggregation. The matching cost C (p, q) computes a similarity value for all potentially matching pixels in the image-pair utilizing the epipolar geometry.[38] In the computation, regularization is used for a proper reconstruction. For SGM, the matching step is cast into an energy minimization problem, and the discontinuity preserving energy is minimized with equation 3.

$$E(D) = \sum_{p} C(p, D_p) + \sum_{p,q \in N_p} V(D_p, D_q) \tag{3}$$

with

$$V(d, d') = \begin{cases} 0 & if\ d = d' \\ P1\ if\ |d - d'| = 1 \\ P2 & otherwise \end{cases} \tag{4}$$

The matching cost function C takes a disparity map D that encodes the correspondences for each pixel location p in the first image and the other image. The pair wise term $V(p, q)$ penalises disparity changes in the neighborhood N_p of each pixel location. The penalty $P1$ is added for all disparity changes equal to one pixel. At more significant discontinuities (disparity change > 1 pixel), a fixed cost $P2$ is added. This cost function favors similar or slightly changing disparities between neighborhood pixels, thus stabilizing the matching in image areas with weak contrast and allowing large disparity jumps in areas with high contrast.[39] SGM can derive a suitable disparity for each pixel with spatial smoothness considered while spending reasonable runtime

proportional to the reconstructed volume.[40] Adjusting the time, finding the optimal disparity image is crucial for SGM.

Minimizing $E(D)$ in a two-dimensional manner is a very costly process that's why SGM simplified the optimization by traversing one-dimensional paths and ensures the constraints with respect to these explicit directions. This approach requires a second phase known as cost aggregation.[41] In SGM, the minimization is realized by cost aggregation along a path in direction r as in equation 5.

$$\begin{aligned} L_r(p,d) = C(p,d) \\ +min_{d'}\big(L_r(p-r,d') + V(d,d')\big) \\ -min\big(L_r(p-r)\big) \end{aligned} \tag{5}$$

Adding the costs L of paths in all directions r provides the aggregated cost S as given equation 6.

$$S(p,d) = \sum_r L_r(p,d) \tag{6}$$

If a pixel adapts its disparity to the neighborhood, it may be assigned a lower cost for one direction. However, at least eight (8) paths in the horizontal, vertical both diagonal orientations have to be introduced. If the number of paths is increased to 16, the streaking artefacts are reduced, and better results are achieved. The disparity map D is computed by identification of the minimum aggregated cost S for each pixel p in the first image.

The disparity image was transformed into the DSM with the desired projection and grid interval. In SGM DSM, the sharp object boundaries are achieved with the advantage of the sudden changeable depth at any pixel against area-based matching methods such as Cross-Correlation, Fourier, Mutual Information and Optimization.[22,42] During the processes, the main disadvantage of SGM was the longer calculation time against LSM.

SGM DSM was generated in PCI Geomatica software, and reprojected epipolar stereo pair with a common orientation was required. Using epipolar stereopair, the possibility of incorrect matches was reduced. For increasing the accuracy of epipolar line alignment, very significant for generated DSM's quality, we used epipolar tracking method. This method enables to follow changes in the epipolar line over the stereo pair and automatically compensates small gradual errors.

For the generation of WV-4 LSM DSM, DPLX and RAPORIO software, developed by Leibniz University Hannover, were used. The visualizations of the DSMs have been done with programs LISA and Surfer.

3.3 DSM Validation

The model comprehensively analyzed generated WV-4 DSMs by LSM and SGM to model visual and statistical approaches with a reference airborne laser scanning (ALS) DSM. The reference ALS DSM has 1 m grid spacing and covers the study area without any remarkable systematic errors, distortions or gaps. Its geolocation accuracy was thoroughly investigated by comparing terrestrial laser scanning (TLS) data in [43], and absolute horizontal and vertical geolocation standard deviations were presented as ≤ 6 cm and ≤ 25 cm, respectively. The ALS point cloud of the reference DSM was produced from 16 points/m^2 with flights organized by Istanbul Metropolitan Municipality in 2013. The DSM is still actual in the study area because there is no significant change.

In DSM validation, the first stage should be the pre-corrections of the tested DSM and the reference. The coordinate system and horizontal and vertical datum have to be the same. In the study, Universal Transverse Mercator (UTM) 35° (Istanbul) was used with World Geodetic System 1984 (WGS84) as the horizontal datum, and orthometric heights were preferred as the

vertical datum. The other case is the 100% horizontal overlap of the tested DSM and the reference DSM for correct vertical accuracy validation. A horizontal offset between the models influences the height corresponding to equation (7) and causes misleading vertical accuracy results. Horizontal offsets were eliminated by shifting the analysed DSMs to the reference DSM corresponding to formula 7. In equation 7, DZ is the vertical discrepancy, DL is the horizontal offset and α is the terrain slope.

$$DL = \frac{DZ}{tan(\alpha)} \tag{7}$$

After all pre-corrections, the vertical absolute and relative geolocation accuracies of WV-4 LSM and SGM DSMs were analysed in three land cover classes as open, built-up, and forest, shown in Figure 4. The built-up areas are dominant in the study area, with 50.14%. The land cover was classified by manual vectorization from VHR WV-4 ortho-image in NetCAD software.

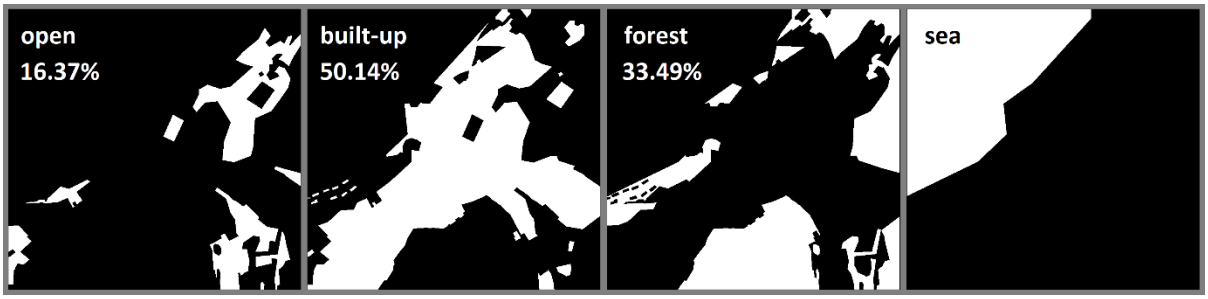


Fig. 4 Land cover classes in the study area

In absolute vertical geolocation accuracy analysis, the standard deviation (STD_Z) and the normalised median absolute deviation ($NMAD_Z$) of height differences between WV-4 LSM and SGM DSMs and the reference was estimated by equation (8) and (9). $NMAD_Z$ is the derivative of median absolute deviation (MAD_Z) of pixel height discrepancies between analysed DSM and

the reference and very effective for the description of significant outliers. However, it is not as sensitive as STD_Z on large individual height differences. MAD_Z is normalised by multiplication factor of 1.4826 to $NMAD_Z$ to reach the same probability of the STD_Z with 68.27% from 50% probability of MAD_Z. If the height differences are normal distributed, which means a symmetric distribution with a single peak (mode) in height distribution histogram, STD_Z and $NMAD_Z$ will be identical. In the forest and built-up areas, large discrepancies occur with the influence of the sudden height changes. In those cases, the normal distribution based on $NMAD_Z$ fits better to the frequency distribution of height differences in comparison with normal distribution based on STD_Z.

$$STD_Z = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\Delta Z_i - \mu)^2} \quad (8)$$

$$NMAD_Z = 1.482615 \times m_i[|\Delta Z_i - m_{\Delta h}|] \quad (9)$$

Where n is the number of height differences between analysed DSM and the reference; ΔZ_i is the individual height differences (errors) between the analysed and reference DSMs $i = 1, \dots\dots, n$. μ is the bias (arithmetic mean of the height differences), m_i is the median of the univariate data set of height discrepancies $i = 1, \dots\dots, n$ and $m_{\Delta h}$ is the median of absolute values of the height errors. Thus, the $NMAD_Z$ is proportional to the median of the absolute differences between errors and the median error.[44]

Besides STD_Z and $NMAD_Z$, skewness and kurtosis coefficients were analyzed in height distribution histograms. In a normal distribution, skewness and kurtosis coefficients should be

around zero. Skewness indicates the lack of symmetry in the height distribution histogram. The negative skew coefficient refers to the pixel height difference distribution enlargement at the left side of the histogram peak, where the positive refers to the right side. Kurtosis is effective in the vertical direction of the height distribution histogram. In positive kurtosis, the histogram peak will be at the upwards of normal distribution peak and indicates that the height differences between analysed DSM and the reference mostly concentrate around one value while the negative kurtosis places in the downward of the normal distribution peak. For the calculation of skewness and kurtosis, equations 10 and 11 were used for an univariate data set of height discrepancies (ΔZ_1, ΔZ_2, ΔZ_3, ….. , ΔZ_n) where μ is the arithmetic mean and n is the number of pixels.

$$skewness = \frac{\sum_{i=1}^{n}(\Delta Z_i - \mu)^3/n}{{STD_Z}^3} \tag{10}$$

$$kurtosis = \frac{\sum_{i=1}^{n}(\Delta Z_i - \mu)^4/n}{{STD_Z}^4} \tag{11}$$

For visual interpretation of height errors between generated WV-4 DSMs and the reference ALS DSM by detecting erroneous regions (if available) and influence of land cover classes, colour-coded height error maps (HEMs) were generated. HEMs are the differential models of analysed and reference DSM and basically calculated with equation 12. Empirically, the height error thresholds are used as ±15 m (~ three times the standard deviation), the most suitable value to recognize the performances of LSM and SGM DSMs. Besides HEMs, an arbitrary profile was detected in a 350 m section, and the vertical coherence of generated DSMs with reference was analysed in X-direction.

$$HEM = DSM_{reference} - DSM_{evaluated} \quad (12)$$

For demonstrating the morphologic performance of generated WV-4 DSMs, contour maps with 10 m equidistance were generated, and the relative standard deviation of height differences $RSTD_Z$ was estimated. The $RSTD_Z$ identifies the interior coherence between neighbour pixels on generated DSMs (equation 13). In this study, for each position, the distance ranges up to point spacing times 10 is used. n_Z indicates the number of point combinations in the distance group. A factor of 2 for multiplication with n_Z is used for normalization of the $RSTD_Z$ to STD_Z. If the height differences of the neighboured points are independent, corresponding to the error propagation, the $RSTD_Z$ would be the square root of 2.0 larger than STD_Z, which is respected by $2 \times n_Z$.

$$RSTD_Z = \sqrt{\frac{\sum(\Delta Z_i - \Delta Z_j)^2}{2 \times n_Z}} \quad \text{Computed for distance groups} \quad (13)$$

4 Results

As mentioned in the introduction, the effect of image matching algorithms on the final DSMs is crucial and a hot research topic for the scientific community. Accordingly, the visual and statistical comparison results of LSM and SGM on WV-4 DSM quality is given in this section. Figure 5 showed the reference ALS DSM and generated LSM and SGM WV-4 DSMs in the study area beside an instance region of interest (ROI) for better interpretation. As aforementioned, the study area is dominant with buildings as 50%, and the superiority of SGM in the growing region is obvious compared to LSM in Figure 5. The shapes of most buildings in

SGM are very close to their real situation given in the ALS reference. Besides, linear details as roads are depicted with high performance in SGM DSM.

In contrast to SGM, the area-based LSM has been generated with sub-windows of 10 × 10 pixels. Corresponding to such sub-windows, it cannot describe buildings as the SGM; it is showing small buildings as small hills. It requires contrast in the sub-windows for matching. As a consequence, the description potential of building shapes in LSM is not at a sufficient level, considering the spatial resolution of WV-4 data. The linear details as roads are available but shown with a noisy structure. Besides, some problems are also stand out in open areas. The visual results in Fig. 5 are in parallel with [22].

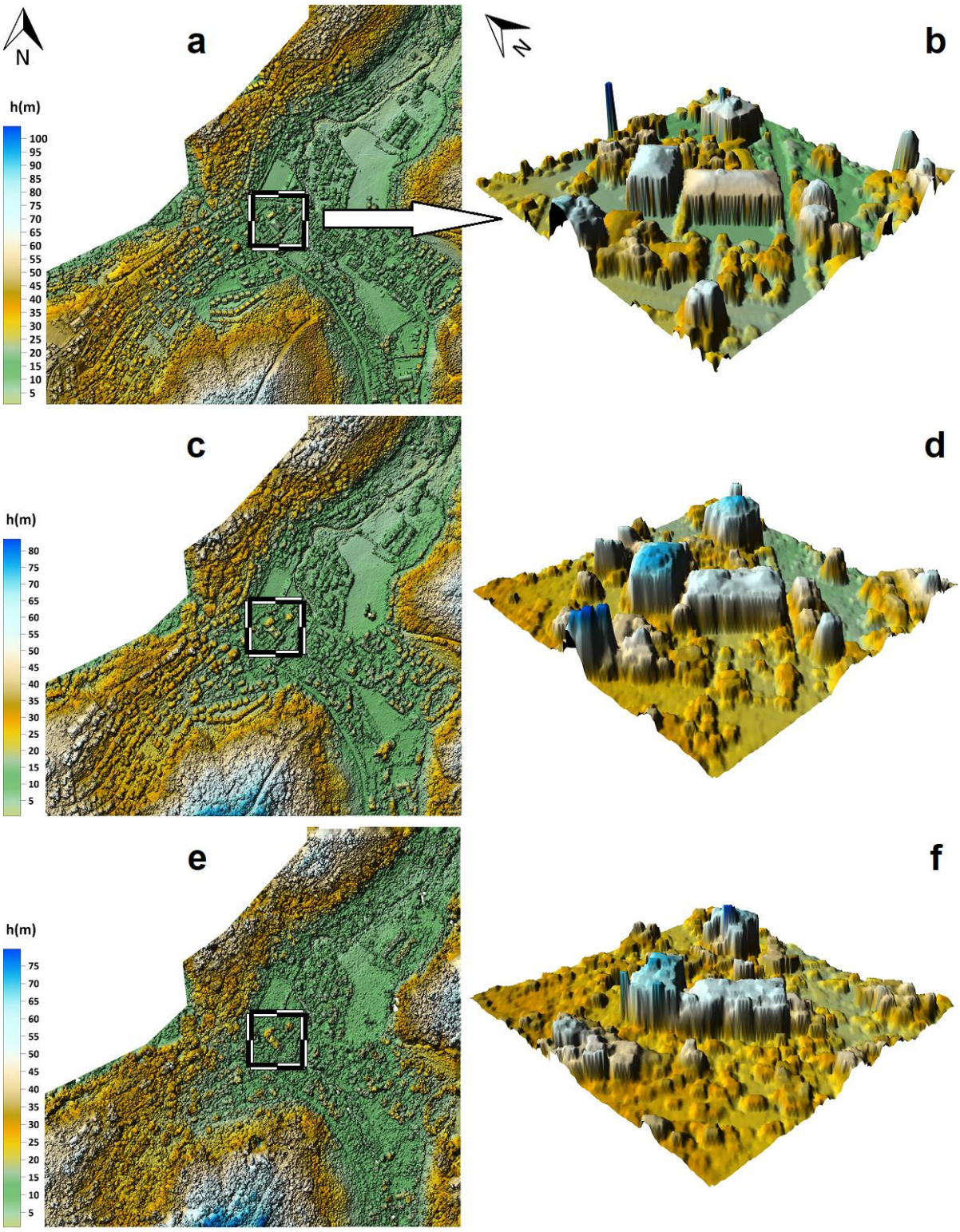


Fig. 5 (a) Reference ALS DSM, (b) Reference ALS DSM ROI, (c) WV-4 SGM DSM, (d) WV-4 SGM DSM ROI, (e) WV-4 LSM DSM, (f) WV-4 LSM DSM ROI

The absolute geolocation errors of LSM and SGM DSMs, calculated by the model to model comparison with the reference ALS, are shown in Table 3 and Table 4. The detected and eliminated horizontal geolocation errors of the LSM and SGM DSMs are very close for both in X with -0.12 m respectively -0.18 m and in Y with 0.11 m respectively 0.14 m. The absolute vertical geolocation accuracies in Table 4 were estimated for each land cover class separately regarding the effect of terrain slope. Besides STD_Z and $NMAD_Z$, systematic vertical bias, skewness and kurtosis values are given in this table as well. In the calculations, the points which have >15 m height difference with the reference are excluded. Also, the percentages of these points are shown in Table 4. In absolute vertical accuracy assessment, the results in open areas are accepted as the most reliable by the scientific community because it does not change suddenly.[45] In Table 4, the performance of both LSM and SGM DSMs in open areas are higher than for other classes, as expected. In open areas, while the standard deviation of LSM DSM is ±2.08 m and the $NMAD_Z$ is ±1.96 m, the SGM has an STD_Z of ±0.98 m and an $NMAD_Z$ of ±0.60 m. In flat terrain, without the influence of slope, the STD_Z reaches ±1.46 m for LSM and ±0.51 m for SGM DSMs, and the $NMAD_Z$ reaches ±1.42 m for LSM and ±0.24 m for SGM DSMs. In open areas, kurtosis values also reach maximum levels because of the upward peak of the frequency distribution against normal distribution.

In the built-up area, both DSMs have declined in vertical absolute geolocation accuracy because of sudden height changes and the effect of buildings' vertical structure, which complicates the correct 3D determination from space depending on sensor viewing geometries with a height to base relation of 1:1.2. A smaller angle of convergence would have an advantage for matching in built-up areas. In most cases, a smaller angle of convergence generates a higher number of matching points in built-up areas, and a large angle of convergence causes a notably larger

amount of missing data. Considering DSM completeness, occlusions usually caused by larger convergence angles on the ground.[46,47] With a standard deviation of ±1.88 m and ±1.12 m, $NMAD_Z$, SGM DSM's performance is much better than LSM in built-up areas. Also, most of the accuracy values for the LSM DSM are larger than SGM DSM in the forest.

Table 3 Estimated and eliminated horizontal geolocation offsets (errors)

Reference DSM	Area	Shifted DSM	ΔX (m)	ΔY (m)
ALS (1m)	Kilyos	WV-4 LSM (1m)	-0.12	0.11
		WV-4 SGM (1m)	-0.18	0.14

Table 4 Vertical geolocation accuracies of WV-4 LSM and SGM DSMs considering different land cover classes and effect of terrain slope (α= slope)

Ref. DSM	Tested DSM	Land Cover Class	Bias (m)	STD_Z (bias eliminated) (m)		$NMAD_Z$ (bias eliminated) (m)		KURTOSIS (m)		SKEWNESS (m)		Excluded Points (%) (>15m)
				STD_Z	STD_Z α < 6°	$NMAD_Z$	$NMAD_Z$ α < 6°		α < 6°		α < 6°	
ALS (1m)	WV-4 LSM (1m)	Whole area 100%	-0.05	3.07+0.41 ×tan(α)	2.08	2.29+0.60 ×tan(α)	1.62	2.74	5.91	0.23	-1.14	0.53
		Open area 16.37%	-0.11	2.08+0.20 ×tan(α)	1.46	1.96+0.20 ×tan(α)	1.42	4.27	0.76	0.77	-0.23	0.29
		Built-up Area 50.14%	0.13	3.39	2.34	2.47+0.19 ×tan(α)	1.58	2.72	5.67	0.15	-1.26	0.67
		Forest area 33.49%	-0.01	3.25+0.69 ×tan(α)	2.52	2.63+0.88 ×tan(α)	1.98	1.05	4.02	0.19	-1.07	1.21
	WV-4 SGM (1m)	Whole area 100%	0.30	1.62+0.42 ×tan(α)	1.20	0.98+0.56 ×tan(α)	0.36	7.02	17.10	-0.73	-2.77	0.06
		Open area 16.37%	0.73	0.98+0.29 ×tan(α)	0.51	0.60+0.38 ×tan(α)	0.24	33.92	62.99	0.42	-4.48	0.26
		Built-up Area 50.14%	0.31	1.88+0.33 ×tan(α)	1.50	1.12+0.41 ×tan(α)	0.48	6.65	13.17	-0.81	-2.78	0.22
		Forest area 33.49%	-0.13	1.53+0.39 ×tan(α)	1.22	1.17+0.49 ×tan(α)	0.49	4.05	6.14	-0.24	-1.16	0.28

In Figure 6, histograms of height differences between generated DSMs and the reference ALS are given for open, built-up and forest areas separately. In addition to the frequency distribution

of height differences (FD_Z) also the normal distributions based on the STD_Z and $NMAD_Z$ are presented. The mentioned problem of LSM DSM against the reference DSM in open areas can be seen in the histogram. The frequency distribution of height errors has double-peak around ±1.5 m. A similar case is valid for LSM DSM in the forest. At least partially, this situation can be explained by the different description of the forest surface by laser scanning and optical image matching. In all land classes, the better performance of SGM is precise.

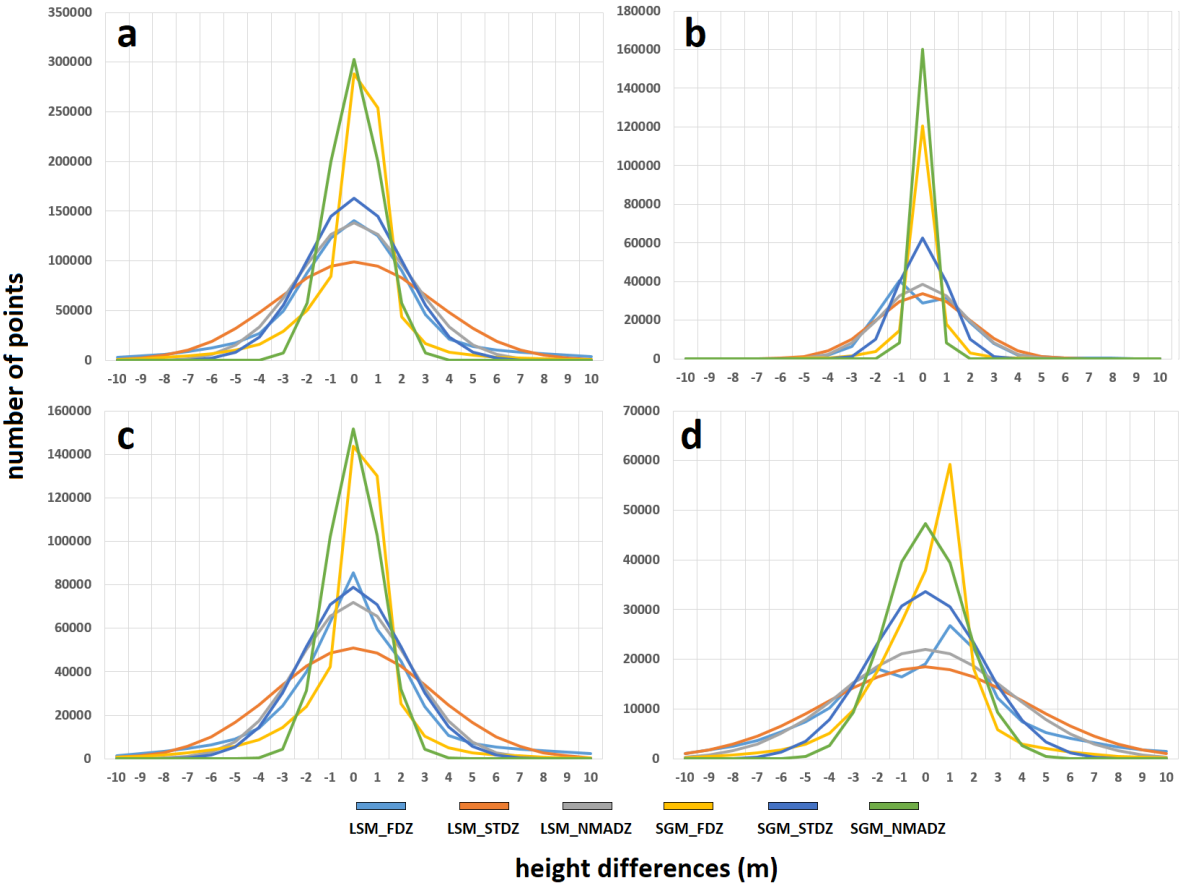


Fig. 6 Distribution of height differences between generated DSMs and reference ALS: (a) whole area, (b) open area, (c) built-up area, (d) forest area

Figure 7 shows the histograms of height differences for flat areas with an inclination below 5.7° (tangent < 0.1). In flat areas, the peaks of FD_Z and $NMAD_Z$ are much higher as for the normal distribution based on STD_Z for all of the land cover classes in SGM DSM. The normal distribution based on STD_Z is not as high in the center due to a higher percentage of larger

discrepancies. In an open and flat area, the normal distribution based on STD_Z fits better to the frequency distribution in SGM. Except for forest, the modes of $NMAD_Z$ are higher than STD_Z and FD_Z in all land cover classes in SGM. In open and flat areas, the performance of SGM is maximum, coherent with the results in Table 4. On the other hand, LSM has double-peaks in open and forest regions.

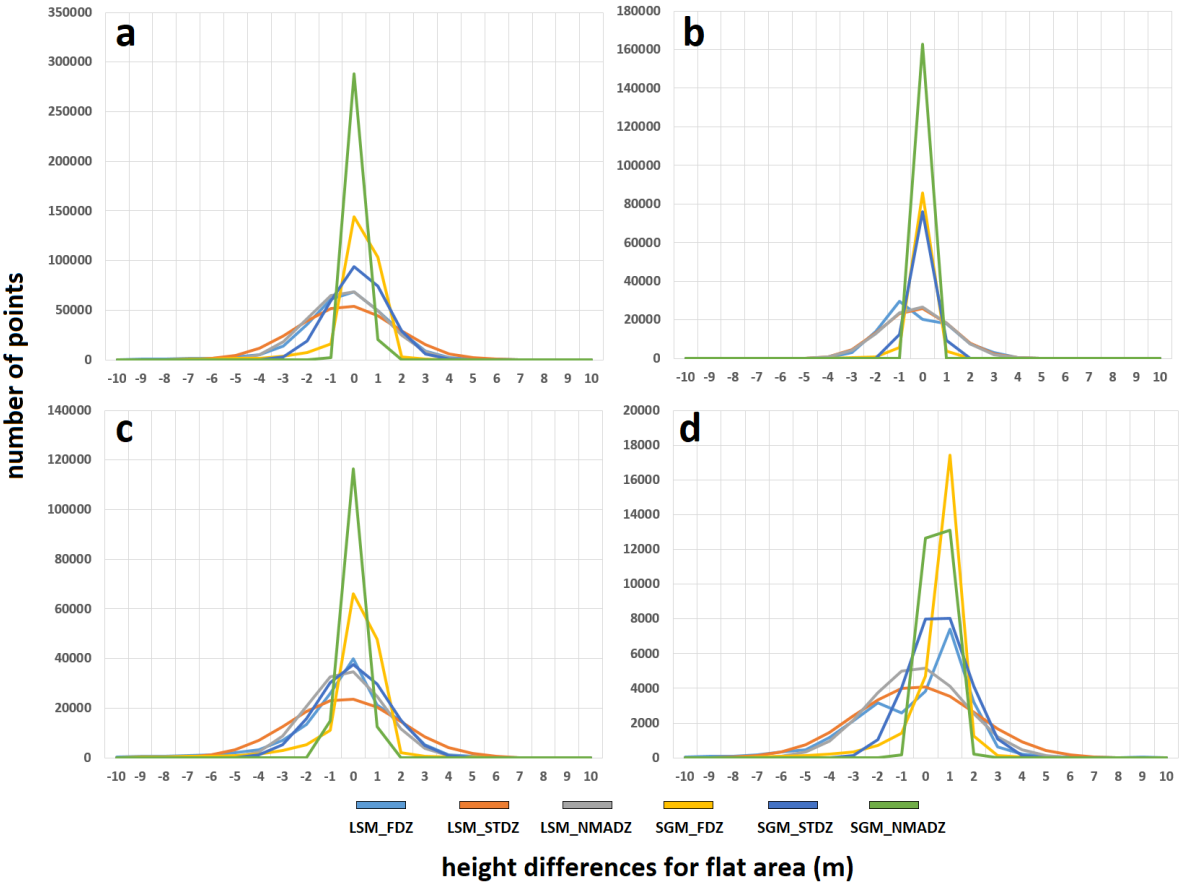


Fig. 7 Distribution of height differences between generated DSMs and reference ALS for uninclined areas: (a) whole area, (b) open area, (c) built-up area, (d) forest area

Figure 8 shows the colour coded HEMs of LSM and SGM DSMs provided with the height differences of corresponding pixels in ALS reference. The error distribution histogram for ±15 m and 20 m height steps are given together with HEMs to interpret the effect of terrain inclination. In the HEMs, the height differences were estimated for all pixels, and light green parts mean maximum coherence and minimum error, where the dark parts mean the lowest coherence and

maximum error. In the HEMs, the topographic description potential differences between LSM and SGM are apparent. The SGM DSM looks more coherent with the reference against LSM, and the light green areas are dominant. On the other hand, LSM DSM has problems in every land classes, and partially the errors reach up to 15 m in built-up and forest areas.

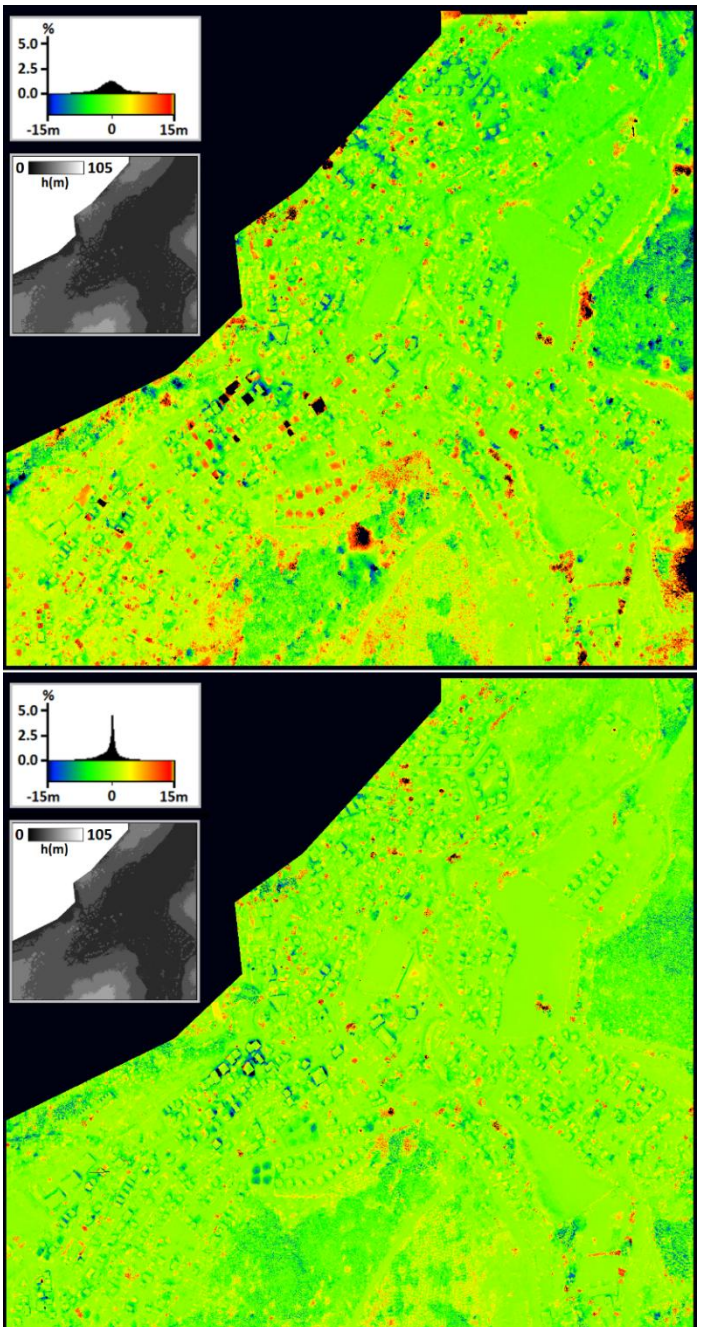


Fig. 8 Generated HEMs overlapped: (a) LSM DSM, (b) SGM DSM

Figure 9 shows the vertical profile coherence in the X direction between LSM and SGM DSMs with the reference throughout 350 m long cross-section. As shown in the figure, different land cover classes are available in the cross-section. In the open area on the left-hand side, the profiles are fitting better together. In the built-up area, it has to be respected that the stereo scene has a height to base relation of 1:1.2. Corresponding to this in the orbit direction, the ground can only be seen stereoscopically if the distance between the buildings exceeds 83% of the building height, and this condition in most cases is not given. For LSM and not only a single-pixel, but the larger part of the used matching window must also be available stereoscopically. This situation explains why SGM DSM looks closer to the ALS profile in the vertical profiles compared to the LSM DSM.

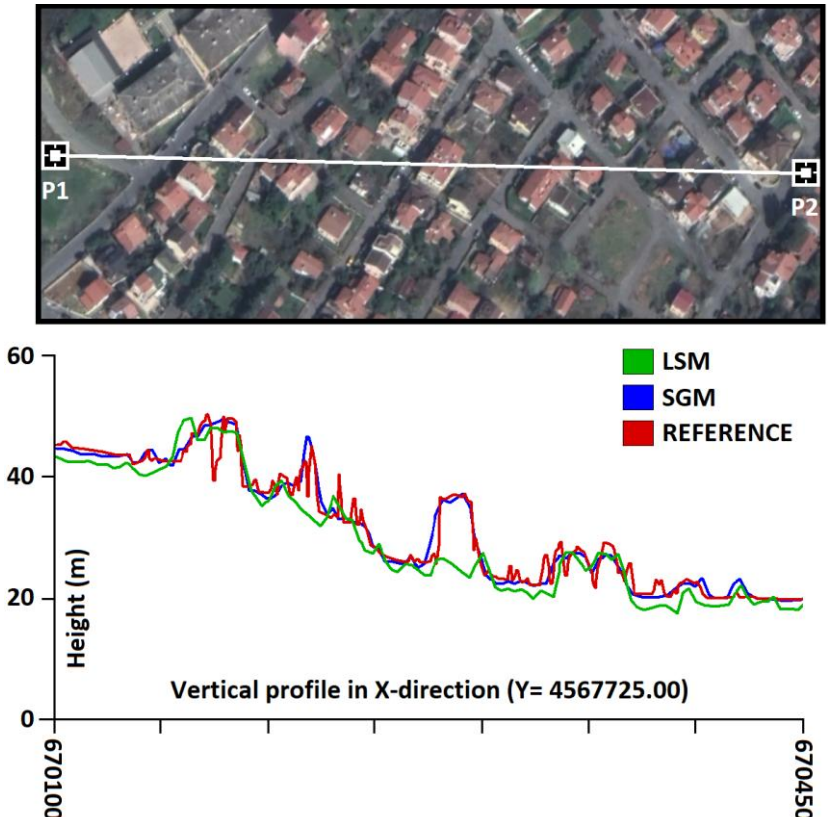


Fig. 9 Vertical profile over 350 m of ALS reference and evaluated DSMs

Contour maps with 10 m equidistance and the relative vertical accuracies, the morphologic performance indicators of generated WV-4 DSMs, are presented in Figure 10 and Table 5. The

contour lines of the DSM cannot be compared with the contour lines of the bare ground. In the DSM, the contour lines are going around most buildings, as the reference clearly shows. The LSM needs more free space around the building and cannot express vertical elements as the SGM. It explains the character of details in Figure 10a. Due to the height to base relation also the SGM has problems in showing the individual buildings. The high noise of a part of the contour lines of the reference is caused by vegetation. The $RSTD_Z$ performance of the SGM DSM pixels is better than the LSM DSM in all land classes, similar to absolute accuracy. When interpreting the relative accuracy, it has to be respected that the definition of the DSM surface determined by ALS (first return) is not the same for matching by optical images.

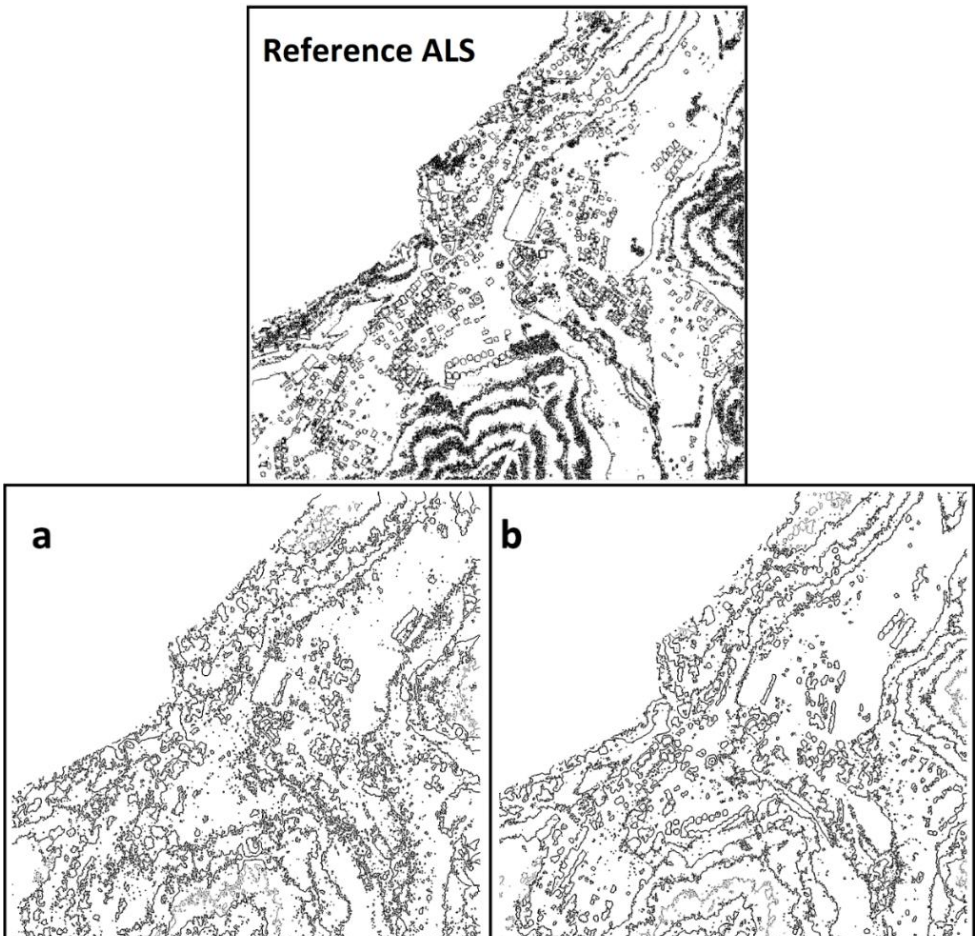


Fig. 10 Morphological case: (a) 10 m interval contour-lines generated from LSM, (b) 10 m interval contour-lines generated from SGM

Table 5 The relative vertical standard deviation of tested DSMs as a function of point distance

Ref. DSM	Distance [10m]	$RSTD_Z$ (m) WV-4 LSM				WV-4 SGM			
		whole	open	built up	forest	whole	open	built up	forest
ALS (1m)	1	1.40	.78	1.47	1.79	1.12	.64	1.19	1.40
	2	1.86	.99	1.98	2.33	1.44	.80	1.56	1.74
	3	2.15	1.11	2.29	2.66	1.62	.88	1.76	1.91
	4	2.34	1.18	2.51	2.87	1.72	.93	1.88	2.01
	5	2.48	1.23	2.66	3.03	1.79	.97	1.96	2.06
	6	2.59	1.26	2.78	3.14	1.83	.99	2.01	2.10
	7	2.67	1.28	2.86	3.23	1.87	1.00	2.05	2.13
	8	2.73	1.29	2.92	3.29	1.89	1.01	2.08	2.16
	9	2.77	1.30	2.97	3.34	1.92	1.02	2.10	2.18
	10	2.80	1.31	3.00	3.38	1.93	1.03	2.12	2.20

5 Discussion of Results

In this section, the absolute vertical geolocation quality of generated WV-4 DSMs is discussed by comparing the results achieved from optical and radar satellites in previous reference studies of our research group in Istanbul. Inaccuracy discussion regarding previous studies shows that the results in open areas are considered because of being the most reliable land cover class with a stable structure independent of the study area.

In [15], a DSM has been generated with WV-2 stereo-pairs in northern Istanbul and comprehensively analysed with the model to model-based approaches using different reference DSMs. The standard deviation of the height based on WV-2 (0.5 m GSD) reached ±1.93 m against an ALS reference DSM in that study. Table 6 shows the standard deviation of height of the optical and radar satellites that our research group achieved in Istanbul approaching in comparisons to ALS and photogrammetric reference DSMs. The photogrammetric reference DSM in Table 6 was produced in 2007-2009 within the scope of a large scale photogrammetry

project of Istanbul Metropolitan Municipality. It has 3 m original grid spacing and ±10 cm absolute vertical accuracy.

Table 6 Absolute vertical geolocation accuracies of different satellites' DSMs

Study Area	Reference DSM	Analysed DSM	GSD (m)	DSM spacing (m)	STD_Z (m)	STD_Z (GSD)	Reference study
Istanbul	ALS	WV-4 SGM	0.30	1	0.51	1.7	This study
	ALS	WV-4 LSM	0.30	1	1.46	4.9	This study
	ALS	WV-2	0.50	1	1.93	3.9	[15]
	ALS	KOMPSAT-3	0.70	5	2.80	4.0	[49]
	Photogrammetry	IKONOS	1	3	3.58	3.6	[51]
	ALS	ALOS AW3D30	2.5	30	4.89	2.0	[48]
	Photogrammetry	SPOT-5 HRS	5	20	4.25	0.9	[50]
	Photogrammetry	ASTER DEM	15	80	6.65	0.4	[50]
	Photogrammetry	ASTER GDEM	15	30	5.33	0.4	[50]
	Photogrammetry	SRTM C-band	20×30	90	3.69	SAR	[50]
	ALS	SRTM X-band	20×30	30	5.36	SAR	[48]
	Photogrammetry	TerraSAR-X HS	1	3	6.10	SAR	[51]
	ALS	Cosmo-SkyMed	1	3	7.33	SAR	[52]
	ALS	Sentinel-1A	5×20	15	5.07	SAR	[48]

In Table 6, it is clear that the standard deviation of height of WV-4 SGM DSM is the highest in compared to other optical and radar satellites' DSMs in Istanbul. However, it also demonstrates that the accuracy is not linear depending on the GSD; it also strongly depends on morphologic details, which are more complicated with higher ground resolution. In the densely built-up area of Istanbul's city, synthetic aperture radar (SAR) has problems with layover and imaging shadows.

5 Conclusion

In this study, the topographic description potential of the latest VHR WV-4 optical satellite was investigated with visual and statistical approaches. The WV-4 DSMs were generated by least

squares and by semi-global matching algorithms separately and evaluated against an ALS reference model. The effects of LSM and SGM algorithms on DSM quality were analysed in detail. The analysis took place in Kilyos district, located in the Istanbul Metropolitan. The test area has a rough topographic structure, and three different land cover classes are open, built-up and forest. The built-up areas are dominant with 50% of the area. To demonstrate the influence of terrain slope on WV-4 DSM quality, the analysis was done separately for the whole area and flat areas (<5.7° inclination).

The statistical results of absolute and relative vertical quality assessment and visual analysis demonstrated that the SGM technique is more accurate and includes more details than LSM for DSM generation with VHR space-borne optical data in urban areas. The horizontal geolocation performances of both DSMs are high and around ±0.3-0.6 pixels. The absolute vertical geolocation accuracy of WV-4 SGM DSM reaches up to ±0.51 m (1.7 pixels) as STD_Z and ±0.24 m (0.8 pixels) as $NMAD_Z$ in open and flat areas where LSM DSM is limited with ±1.46 (4.9 pixels) and ±1.42 m (4.7 pixels) respectively. The absolute vertical accuracy of SGM DSM is the highest value that our research group achieved from an optical space-borne mission. Besides, relative accuracies of the WV-4 SGM DSM are very high, particularly in open areas. Generated height error maps prove that LSM DSM is not as accurate as SGM in all land cover classes and flat topography. According to the HEMs and vertical profiles in 350 m cross-section, WV-4 SGM DSM looks very coherent with the reference ALS DSM which supports the numerical results.

Acknowledgements

We would like to thank Yıldız Technical University for supporting the acquisition of WV-4 stereo imagery in the scope of the Scientific Research Project with ID: 3075. Besides, many thanks to the Istanbul Metropolitan Municipality for reference ALS data.

References

1. A. El Garouani, and A. Alobeid, "Digital Surface Model Generation for 3D City Modeling (Fez, Morocco), 8th National GIS Symposium in Saudi Arabia," in *The Road for Building Saudi Arabia GIS" organized by the high Committee of the GIS at the Eastern Province, Dammam, Saudi Arabia*, April 15–17 (2013).
2. M. Font, D. Amorese, and H.L. Lagarde, "DEM and GIS Analysis of The Stream Gradient Index to Evaluate Effects of Tectonics: The Normandy Intraplate Area (NW France)," *Geomorphology*, 119(3–4), 172– 180 (2010).
3. U.G. Sefercik, et al., "Novel Fusion Approach on Automatic Object Extraction from Spatial Data: Case Study Worldview-2 and TOPO5000," *Geocarto Int.* **33**(10), 1139-1154 (2018b).
4. R. R. Navalgund, V. Jayaraman, and P.S. Roy, "Remote Sensing Applications: An Overview," *Current Sci.* **93**(12), 1747–1766 (2007).
5. U.G. Sefercik, and A. Atesoglu, "Three-Dimensional Forest Stand Height Map Production Utilizing Airborne Laser Scanning Dense Point Clouds and Precise Quality Evaluation," *iForest,* **10**, 491-497 (2017b).
6. D. Tapete, and F. Cigna, "COSMO-SkyMed SAR for Detection and Monitoring of Archaeological and Cultural Heritage Sites," *Remote Sen-Basel,* **11**, 1326 (2019).
7. M. Alkan, et al., "Geometric accuracy and information content of WorldView-1 images," *Optical Eng.* **52**(2), 026201 (2013).

8. D. Poli, et al., “Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pleiades-1A stereo images for 3D information extraction,” *ISPRS J. Photogramm. Remote Sens.*, **100**, 35-47 (2015).
9. F. Hu, X. Gao, M., G.Y. Li, and M. L, “DEM Extraction from Worldview-3 Stereo-Images and Accuracy Evaluation. ISPRS,” in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1: 327-332 (2016).
10. T. Wang, and A. Gruen, “Georeferencing Accuracy Analysis of WorldView-2 and IKONOS Images of Singapore Based on RPFs.,” in *33rd Asian Conference on Remote Sensing (ACRS) 2012*, pp. 1494-1501 (2012).
11. T. Toutin, C.V. Schmitt, and H. Wang, “Impact of no GCP on elevation extraction from WorldView stereo data,” *ISPRS J. Photogramm. Remote Sens.* **72**, 73–79 (2012).
12. M. Crespi, et al., “A New Rigorous Model for High Resolution Satellite Imagery Orientation: Application to EROS A and QuickBird,” *Int. J. Remote Sensing*, **33(8)**, 2321–2354 (2012).
13. D. Poli, K. Wolff, and A. Gruen, “Evaluation of Worldview-1 Stereo Scenes and Related 3D Products. International Archives of Photogrammetry,” in *Remote Sensing and Spatial Information Sciences*, Vol. XXXVIII-1-4-7/W5, Hannover (2008).
14. K. Jacobsen, “DEM Generation from High Resolution Satellite Imagery,*” Fernerkund. Photogramm.* **5**, 483-493 (2013).
15. U.G. Sefercik, et al., “Generation and Validation of High-Resolution DEMs from Worldview-2 Stereo Data,” *Photogramm. Rec.* **28**(144), 362-374 (2013a).
16. O. R. Belfiore, and C. Parente, “Comparison of Different Algorithms to Orthorectify Worldview-2 Satellite Imagery,” *Algorithms*, **9**(4), 67 (2016).
17. M. A. Aguilar, et al., “Quality Assessment of Digital Surface Models Extracted from Worldview-2 And Worldview-3 Stereo Pairs Over Different Land Covers,” *GIScience Remote Sens.*, **56**(1), 109-129 (2019).

18. Z. L, and J. Wang, “Least squares image matching: A comparison of the performance of robust estimators,” in *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-1. 37-44. 10.5194/isprsannals-II-1-37-2014* (2014).
19. H. Hirschmuller, “Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information,” in *Proceedings of IEEE Conference on CVPR, New York, NY* (2005).
20. A. Alobeid, et al., “Building Monitoring with Differential DSMs,” in *ISPRS Hannover Workshop* 2011, *International Archives of Photogrammetry and Remote Sensing, XXXVIII-4(W19)* (2011).
21. L. Zhang, and A. Gruen, “Multi-image matching for DSM generation from IKONOS imagery,” *ISPRS J. Photogramm. Remote Sens.*, **60**(3),195-211 (2006).
22. C. Heipke, Automation of Interior, Relative, and Absolute Orientation, *ISPRS J. Photogramm. Remote Sens.*, **52**,1-19 (1997).
23. G. Buyuksalih, M. Oruc, and K. Jacobsen, “Precise Georeferencing of Rectified High Resolution Space Images,” in *XXth ISPRS Congress, Istanbul, Turkey,* 35: 184-188 (2004).
24. T. Toutin, “Geometric Processing of Remote Sensing Images: Models, Algorithms and Methods,” *Int. J. Remote Sens.* **25**,1893-1924 (2004).
25. C. S. Fraser, and H. B. Hanley, “Bias-compensated RPCs for Sensor Orientation of High-resolution Satellite Imagery,” *Photogramm. Eng. Remote Sens.***71**(8), 909–915 (2005).
26. J. Grodecki, “IKONOS Stereo Feature Extraction - RPC Approach.,” in *ASPRS Annual Conference, St. Louis*, 23–27 April (2001).
27. J. Grodecki, and D. Gene, “Block Adjustment of High-Resolution Satellite Images Described by Rational Polynomials,” *Photogramm. Eng. Remote Sens.* **69**(1), 59–68 (2003).
28. I. Najafi, and M. Momeni, “Precision Evaluation of RPCs in HRSI Orientation of Geoeye-1,” in *5th Symposium on Advances and Science & Technology, Mashhad, Iran*, 12–17 May (2011).
29. C. S. Fraser, and H. B. Hanley, “Bias Compensation in Rational Functions for IKONOS Satellite Imagery,” *Photogramm. Eng. Remote Sens.* **69**(1), 53–57 (2003).

30. K. Di, R. Ma, and R. Li, “Rational Functions and Potential for Rigorous Sensor Model Recovery,” *Photogramm. Eng. Remote Sens.* **69**(1), 33–41 (2003).
31. H. Hirschmüller, and T. Bucher, “Evaluation of Digital Surface Models by Semi-Global Matching,” in *DGPF, Vienna, Austria* (2010).
32. F. Bethmann, and T. Luhmann, “Least-Squares Matching with Advanced Geometric Transformation Models,” in *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII(5) (2010).
33. W. Foerstner, “On the Geometric Precision of Digital Correlation,” in *International Archives of Photogrammetry and Remote Sensing*, 24(3): 176-189 (1982).
34. F. Ackermann, “Digital Image Correlation: Performance and Potential in Photogrammetry,” *Photogramm. Record*, **11**(64), 429-439 (1984).
35. A. Pertl, Digital Image Correlation with the Analytical Plotter Planicomp C-100. *International Archives of Photogrammetry and Remote Sensing*, 25 (3B): 874-882 (1984).
36. A. Grun, “Adaptive Least-Squares Correlation – A Powerful Image Matching Technique,” *South African Journal Photogramm. Remote Sens. Cartography*, **14**(3), 175-187 (1985).
37. H. Hirschmüller, “Stereo Processing by Semi-Global Matching and Mutual Information,” in *PAMI*, 30, 328–341 (2008).
38. H. Hirschmüller, and D. Scharstein, “Evaluation of Stereo Matching Costs on Image with Radiometric Differences,” *IEEE Trans. Pattern Analysis Machine Intellig.* **31**(9), 1582–1599 (2009).
39. P. D’angelo, “Improving Semi-Global Matching: Cost Aggregation and Confidence Measure. ISPRS - International Archives of the Photogrammetry,” *Remote Sens.Spatial Inform. Sci.* XLI-B1: 299–304 (2016).
40. P. D’angelo, and P. Reinartz, “Semi-global Matching Results On the ISPRS Stereo Matching Benchmark,” in *ISPRS International Archives of Photogrammetry and Remote Sensing, Spatial Information Science, XXXVIII-4/W19,* 79–84 (2011).

41. M. Michael, et al., “Real-time Stereo Vision: Optimizing Semi-Global Matching,” In *IEEE Intelligent Vehicles Symposium (V)* (2013).

42. J. Joglekar, and S. S. Gedam, “Area Based Image Matching Methods - A Survey,” *Int. J.Emerg. Technology Adv. Eng.* **2**, 130– 136 (2012).

43. U. G. Sefercik, et al., “Point-based and Model-Based Geolocation Analysis of Airborne Laser Scanning Data,” *Opt. Eng.* **56**(1), 013101 (2017a).

44. J. Höhle, and M. Höhle, “Accuracy assessment of digital elevation models by means of robust statistical methods,” *ISPRS J. Photogramm. Remote Sens.*, **64**(4), 398–406 (2009).

45. U. G. Sefercik, et al., “Area-based Quality Control of Airborne Laser Scanning 3D Models for Different Land Classes Using Terrestrial Laser Scanning: Sample Survey in Houston, USA,” *Int. J. Remote Sens.* **36**(23), 5916-5934 (2015).

46. L. Piermattei, et al., “Impact of the Acquisition Geometry of Very High-Resolution Pléiades Imagery on the Accuracy of Canopy Height Models over Forested Alpine Regions,” *Remote Sensing*, **10**(10), 1542 (2018).

47. A. Loghin, J. Otepka-Schremmer, and N. Pfeifer, “Potential of Pléiades and WorldView-3 Tri-Stereo DSMs to Represent Heights of Small Isolated Objects,” *Sensors*, **20**(9), 2695.

48. U. G. Sefercik, et al., “Validation of Sentinel-1A And AW3D30 DSMs for the Metropolitan Area of Istanbul, Turkey,” *J. Photogramm. Remote Sens. Geoinf. Science.* **86**(3-4), 141-155 (2018a).

49. U. G. Sefercik, et al., “ DSM Quality of Korean Satellite KOMPSAT-3 in Comparison to AW3D30 and Sentinel-1A in Respect of Airborne Laser Scanning,” *KSCE J. Civil Eng.* **23**(7), 3162-3173 (2019).

50. N. Yastikli, U. G. Sefercik, and F. Esirtgen, “Quantitative Assessment of Remotely Sensed Global Surface Models Using Various Land Classes Produced from Landsat Data in Istanbul,” *Chinese Geogr. Sci.* **24**(3), 307-316 (2014).

51. U. G. Sefercik, and M. Ozendi, “Comprehensive Comparison of VHR 3D Spatial Data Acquired From IKONOS and TerraSAR-X Imagery,” *Adv. Space Res.* **52**(9), 1655-1667 (2013b).

52. U. G. Sefercik, and N. Yastikli, "ALS Model-Based Comparison of Cosmo-SkyMed and TerraSAR-X HS DSMs on Varied Land Forms," *J. Spat. Sci.* **61**(1), 119-131 (2016).

First Author is a professor at the Gebze Technical University. He received his BS and MS degrees in Geomatics from the Zonguldak Bülent Ecevit University 2005 and 2010 respectively. He finished his postdoctoral at the National Center for Airborne Laser Mapping (NCALM), University of Houston, USA, in 2014. He is the author of more than 20 journal papers. His research interests are processing, generation, and quality validation of airborne and spaceborne remote sensing data and products.

Second Author is a professor at the Yıldız Technical University. He received his BS and MS degrees in Geomatics from the Karadeniz Technical University 1997 and 2005 respectively. He is the author of more than 25 journal papers and has written one book chapter. His research interests are database management, geographical information systems, remote sensing, object extraction, national spatial data infrastructure, e-municipality, and e-government.

Third Author received his PhD in photogrammetry from Leibniz University Hannover, Germany. His main research area is numerical photogrammetry, especially the use of space imagery.

Fourth Author received his MsC from Zonguldak Bulent Ecevit University at 2017. He is an research assistant at Zonguldak Bulent Ecevit University Geomatics Department. His doctoral studies are still continuing at Zonguldak Bulent Ecevit University.

Fifth Author received his PhD from the Department of Geography and Topographic Science, University of Glasgow, UK. He is an associate professor in photogrammetry at Private Sector, Istanbul, Turkey. His research direction is the full range of photogrammetry and remote sensing.