Sampling Designs for Selecting Points from Digital Surface Models to Create Digital Terrain Models

Philip Fayad
Cyprus University of Technology
Department of Civil Engineering and Geomatics
School of Engineering and Technology
30 Archbishop Kyprianou Str., Limassol, Cyprus
pk.fayad@edu.cut.ac.cy

Apostolos Papakonstantinou
Cyprus University of Technology
Department of Civil Engineering and Geomatics
School of Engineering and Technology
2-8 Saripolou Str., Achilleos II Bldg., 3nd floor, Limassol, Cyprus
a.papakonstantinou@cut.ac.cy

Phaedon Kyriakidis
Cyprus University of Technology
Department of Civil Engineering and Geomatics
School of Engineering and Technology
2-8 Saripolou Str., Achilleos I Bldg., 2nd floor, Limassol, Cyprus
phaedon.kyriakidis@cut.ac.cy

Abstract

Recent advances in Unmanned Aerial Vehicle (UAV), in photogrammetric software and the miniaturisation of sensors lead topographers to embed the Unmanned Aerial System - Structure form Motion (UAS-SfM) pipeline to field work for earth measurements and DTM creation. This study attempts to develop a tool and examine a methodology to produce Digital Terrain Model (DTM) from a Digital Surface Model (DSM) created following UAS-SfM pipeline, while investigating the effect of known elevation point's spatial distribution during interpolation. The proposed methodology aims to automate the DTM production process and minimize the production time, while maintaining a high-quality result and a low-cost approach. A geo-processing tool named "RS_Sampling Tool" is developed in order to extract known height points updating altitude information from the DSM using two sampling design schemas, random and stratified. The Inversed Distance Weighting (IDW) interpolation method was selected to create the DTM's from six scenarios, in which the elevation points are differentiating in quantity and spatial distribution. The DSM used was created from a UAS flight realised on 25th of July 2017 to map Vrisa settlement on the island of Lesvos in Greece after a catastrophic earthquake that took place. The main conclusions of this study include: a) known altitude height points should be derived from bare ground areas of the DSM, b) the more points being used from the DSM the more accurate the final DTM is and c) a random sample distribution is much more likely to yield a DTM with low accuracy. However, a random based distribution can yield a high accuracy DTM, by including high-precision GCP's.

Keywords: Sampling Design; DTM; DSM; interpolation; methodology; elevation points.

1 Introduction

Regular and irregular grid, contouring and the sectional diagram are the common graphic representations (Gold, 2005). Digital Terrain Models (DTM's) are produced using elevation sampling points distributed in space, in a way that is representative of the Area of Interest (AoI) and they can be measured using various altitude recording devices like GPS, Lidar, etc. (Heywood et al., 2006). In this way it is possible to apply spatial interpolation methods and generate elevation values throughout the whole AoI, creating a DTM, a Digital Elevation Model (DEM) or a Digital Surface Model (DSM) (Gold, 2005; Watson and Philip, 1985). In general, spatial interpolation is the process of calculating unknown values in space, using some sampling observations (values) and it refers to continuous phenomena throughout the space (Heywood et al., 2006). The terms "DEM" and "DTM" are more specific and refer only to digital altitude representation of the earth's surface without any additional elements, while DSM's refer to surfaces that also contains other physical or anthropogenic elements such as structures, vegetation, etc.

In this study the Inversed Distance Weighting (IDW) technique is used, one of the most popular deterministic interpolation methods, mainly because of its simplicity and efficiency. However, in order to obtain the best results a relatively dense elevation sampling distribution is required (Watson and Philip, 1985).

Due to the very recent advances in the fields of computer vision and photogrammetry and in combination with the improvements in data processing power, orthophoto maps and DSMs can be easily produced by terrestrial and/or aerial high-resolution 2D imagery. A plethora of scientists encourage the use of Unmanned Aerial Vehicle (UAV), according to Adams (2011) and Papakonstantinou (2019), UAVs have been utilized with great potential following the 2009 L'Aquila, 2010 Haiti, 2011 Japan and 2017 Vrisa earthquakes and each event presented different opportunities and lessons that will mould the promising future of UAV usage for imagery collection in disaster management and monitoring. Furthermore, an extensive study was implemented to monitor and to map Vrisa village damage assessment as a post-earthquake process (Soulakellis et al., 2018).

Recent technological advantages make UAV-based photogrammetry highly suitable for surveys in a geo-hazard context, as in a post-earthquake scenario, and its advantages may be summarized as follows: a) safety: no risk for operators; b) possibility to survey inaccessible zones; c) high-resolution photographs; d) speed of survey and elaboration; and e). repeatability and economic convenience (Dominici et al., 2017; Kavroudakis et al., 2018).

This study aims to develop a semi-automated method for selecting points from a DSM in order to produce a very detailed in terms of resolution and earth's surface ground variation DTM. A UAV flew over the Area of Interest (AoI) and the DSM is produced following the Unmanned Aerial System - Structure form Motion (UAS-SfM) pipeline. A total of six sampling scenarios were implemented to derive ground heights from the UAS-DSM. The heights derived from the DSM were used for the creation of new a DTM, one for each scenario. The DTMs produced are compared in order to investigate the effect of known elevation point's spatial distribution during interpolation. More specific in the DTM comparison the spatial distribution of the elevation points is being investigated by varying the quantity and the randomness of the elevation points used in the spatial interpolation process. In the following chapters the proposed methodology is analysed.

2 Materials and Methods

Traditionally the creation of DTM for an area is implemented using spatial interpolation to ground measurements. In this process the calculation of unknown values based on some sampling observations applies to DTM creation.

2.1 Area of Interest

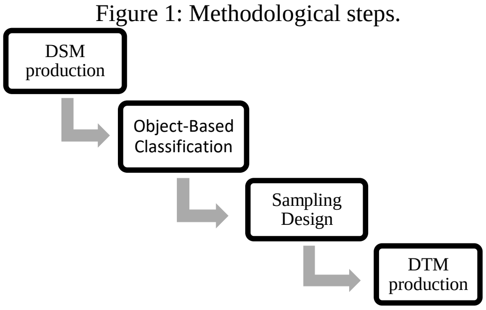
On 12 June 2017 (UTC 12:28:38.26) a magnitude Mw 6.3 earthquake occurred offshore Lesvos Island in SE Aegean Sea, which was widely felt to the citizens of the island. Most substantial damage was reported to the village of Vrisa located to the south-eastern coast of the island. The study area covers all the damaged part of the village, approximately 0.3Km^2.

2.2 Methodology

The workflow of the semi-automated methodology presented in this study (fig.1) consists the following steps:

1. DSM production: a DSM of the AoI is produced following UAS-SfM pipeline,
2. Object-based Classification for the automatic identification and selection of pure ground polygons on the UAS-DSM,
3. Selection of elevation points within the polygons selected from previous step according to the parameters (RS Sampling geoprocessing Tool creation),
4. Spatial interpolation using the IDW approach for the DTM production.

Figure 1: Methodological steps.



On the 25th of July a UAS data acquisition campaign from the University of the Aegean took place for mapping Vrisa Settlement at village spatial scale. In this campaign the following spatial data were collected: a) 148 Ground Control Points (GCP's) using a Real Time Kinematic (RTK) system and b) 1000 very high-resolution aerial images. The UAS flew at low altitude (65 m), capturing high resolution images with 80% overlap and 80% sidelap. Following the UAS-SfM pipeline, the orthophoto map (fig.2) and the DSM (fig.3) created having a spatial resolution of 3 cm and 5 cm, respectively.

Figure 2: The Orthophoto map produced.

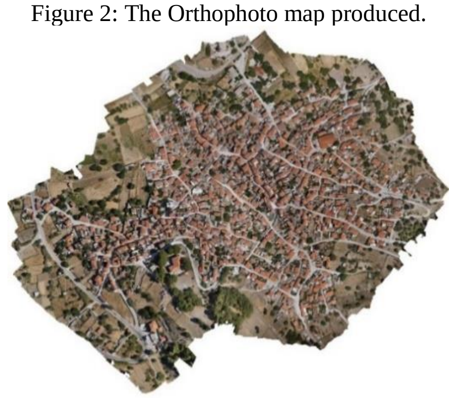
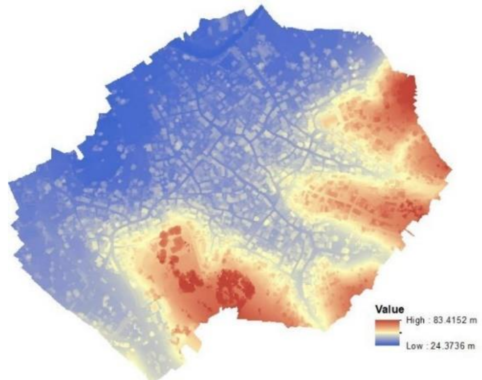


Figure 3: The Digital Surface Model produced.



From the constructed DSM it is crucial to clearly identify bare ground areas in order to select and extract elevation values from these areas. This can be achieved using classification methods. Two are the most important: a) the pixel-based, in which classification is done according to the ground spectrum differences and b) the object-based method, in which classification is relying not only on the spectral characteristics of ground, but also consider geometric and structural information. Many studies, like the one implemented by Chen (2009), had proven the efficiency and the advantages of the Object-based classification method over the traditional pixel-based method especially in reducing errors/noises and "salt and pepper" phenomena. Figure 4 display the process followed in order to define the elevation points in each one of the six scenario.

To define polygons of the bare ground areas in the DSM, the Object-based classification approach was implemented, using the Feature Extraction tool in ENVI 5.0 (Feature Extraction Module User's Guide, 2008). This led to 1250 polygons of various sizes containing a total of 25+ million elevation points. The 626 of the total 1250 polygons were polygons with area less than 10 m^2 (labeled as small areas). The result of this classification process was polygons that contained a large number of elevation points. Following the methodology (fig.1), in order to select the appropriate amount of elevation points needed and reduce the total number so that those can be used in the spatial interpolation process, the next step was to create the different elevation points selection scenarios. That was achieved through the Sampling Design process and using the random or the stratified sampling method. The number of points used was not the same in each of the six scenarios thus, consisting: a) a standard number of 626 points that represent the centroid of the polygons of small areas (areas < $10m^2$) and b) in some scenarios, 55 (out of the total 148) additional high precision (RTK) GCP's. The distribution of the 55 GPS points followed a pattern that cover all the study area. The 35 were selected visually according to the terrain variations of the study area and the remaining 20 according to their height value (where the average height value of the area is exceeded).

In this study, six point sampling scenarios were applied (table 1) using the random and stratified sampling methods. In D, E scenarios the preselected 55 additional GCP's were added to the DTM creation. The sampling parameters used were the following: a) number of points in each polygon, b) minimum allowed distance between points and c) Cell size. The number of points distributed in each polygon was set to 100 at scenario A while at scenarios B and D was depended from the point height range. The minimum allowed distance between points was set as a fixed value (2m) at scenarios A, B

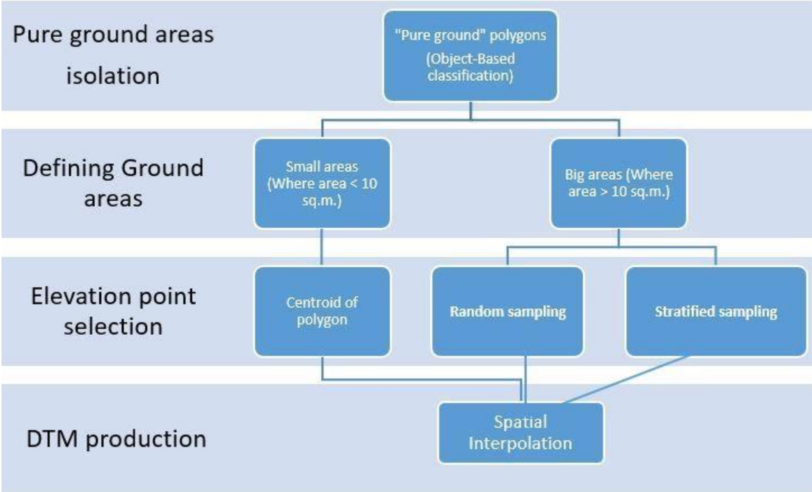
Table 1: Elevation points selection scenarios.

Scenario	Sampling method	Sampling parameters	No of Points used	Interpolation method	Interpolation parameters
A	Random	Number of points in each polygon: fixed = 100, Minimum allowed distance between points: fixed = 2m	3,209 + 626 (of small areas) = 3,835	IDW	p = 2, Search radius Variable, Number of points = 16
B	Random	Number of points in each polygon: Variable = depending on height range, Minimum allowed distance between points: fixed = 2m	1,557 + 626 (of small areas) =2,183	IDW	p = 2, Search radius Variable, Number of points = 16
C	Stratified	Cell height = 3m Cell Width = 3m	7,109 + 626 (of small areas) = 7,735	IDW	p = 2, Search radius Variable, Number of points = 16
D	Random	Number of points in each polygon: Variable = depending on height range, Minimum allowed distance between points: fixed = 2m	1,612 + 55 (GPS) + 626 (of small areas) = 2,293	IDW	p = 2, Search radius Variable, Number of points = 16
E	Stratified	Cell height = 3m Cell Width = 3m	7,109 + 55 (GPS) + 626 (of small areas) = 7,790	IDW	p = 2, Search radius Variable, Number of points = 16
Z	-	All elevation points are used	25,569,387 + 626 (of small areas) = 25,570,013	IDW	p = 2, Search radius Variable, Number of points = 16

and D. At scenarios C and E the parameter cell size was set to 3x3m in height and width respectively. Finally, at scenario Z all elevation points delivered from the DSM were used. In all scenarios, the IDW method was used, with the "Search radius" and "Number of points" as predefined variables having values 2 and 16 respectively.

The "RS Sampling Tool" developed implements the Sampling design calculations and exporting all the resulting elevation points, according to the parameters set. The last step in the proposed methodology is the production of the DTM using the spatial interpolation process.

Figure 4: The elevation points selection process (Sampling Design) for the interpolation and the DTM production.



3 Results

The resulting surfaces produced from the proposed methodology were of small differences (fig.5). All DTMs produced, were having heights variation of 25.503m (scenario C and E) to 73.373m (scenario B). Surface Z (fig.6) was an exception having the lowest height (25.390m) as well as the highest (74.019m) height values among all surfaces. Comparing the six elevation surfaces, scenario Z has resulted to a much rougher surface and that is due to the fact that the total of the 25 million elevation points was used.

The height variation in all the scenarios implemented is depicted in Table 2. Furthermore, a statistical comparison of the DTMs was implemented using the min-max value matrix (table 2), the RMSE table (table 3 and fig.7) and the correlation matrix (table 4). To investigate and visualize the differences between the DTM of scenario Z with all the other scenarios, we used the cut fill tool in ArcGIS (fig.8).

Table 2: Min and Max height value in each scenario.

Scenario	MIN (m)	MAX (m)
A	25.629	73.468
B	25.919	73.373
C	25.503	73.609
D	25.740	73.374
E	25.503	73.609
Z	25.390	74.019

As for the accuracy of the produced DTMs, table 3 depicts the total RMSE errors, which are calculated using the remaining GCP's in each scenario, out of the total 148 GCP's.

Figure 5: The resulting DTM surface in each scenario.

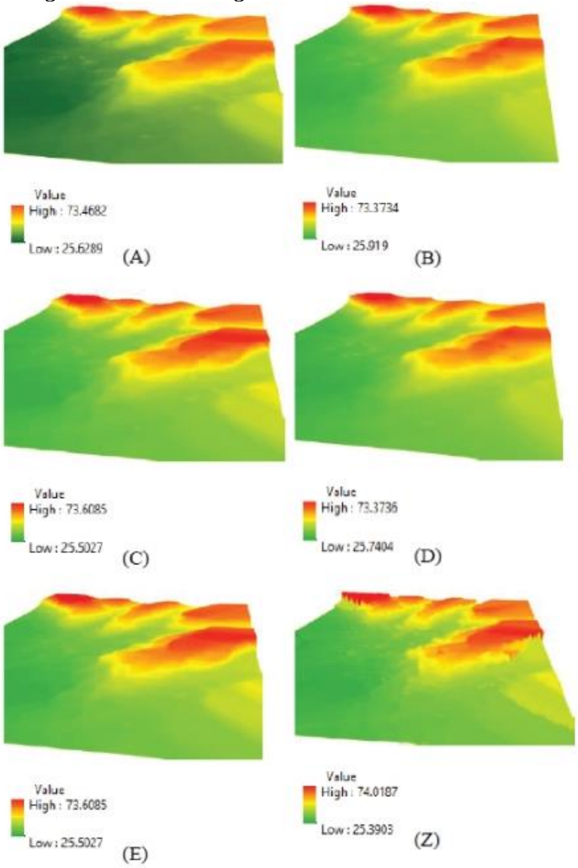


Figure 6: The resulting DTM of scenario Z, where all 25 million elevation points are being used.

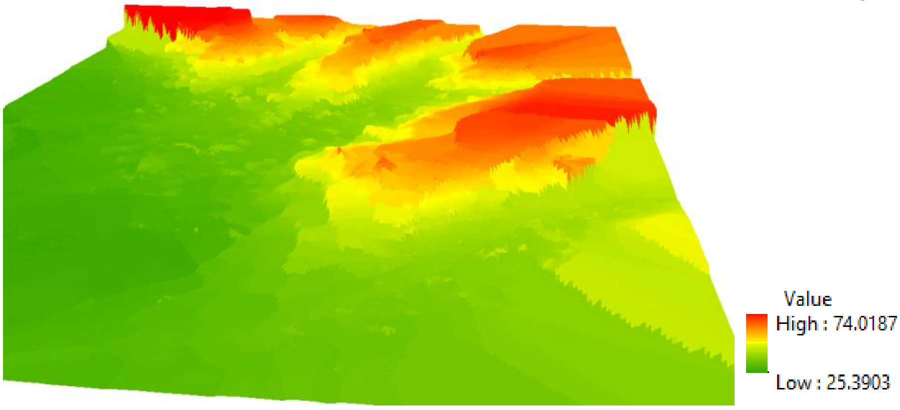


Table 3: The calculated RMSE in each scenario.

Scenario	RMSE (all 148 GCP's used)	RMSE (148 -55 = 93 GCP's used)	Number of Elev. points
A	1.007m	-----	3,835
B	1.108m	-----	2,183
C	0.620m	-----	7,735
D	-----	0.897m	2,293
E	-----	0.590m	7,790
Z	0.377m	-----	25,570,013

Figure 7: The number of elevation points used in the DTM generation in correlation with the RMSE that resulted in each scenario.

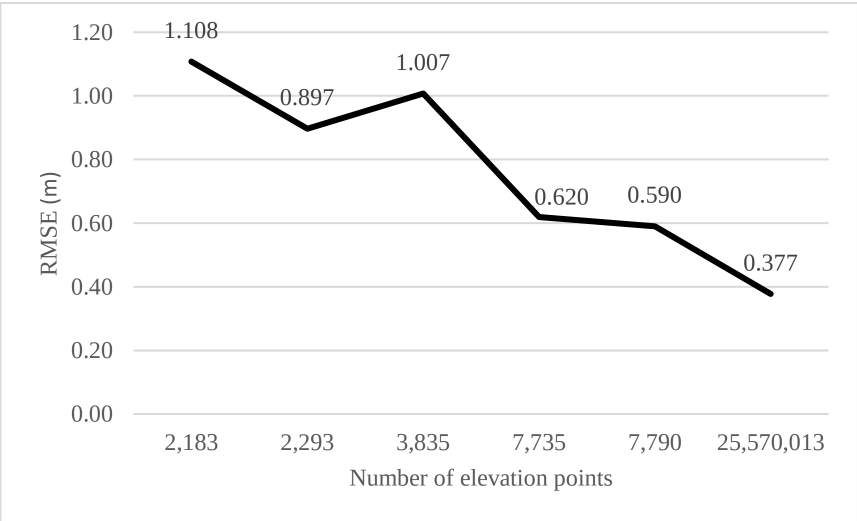
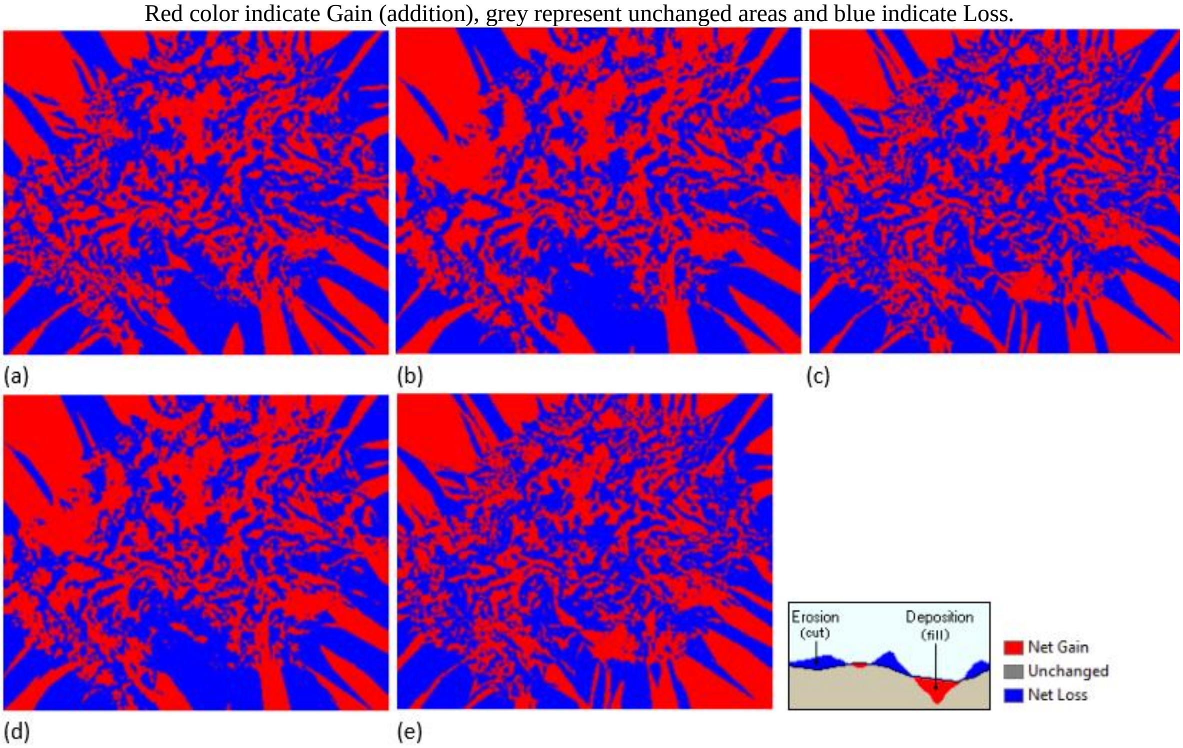


Table 4: Correlation Matrix where values range between 0 and 1. Zero value means absolutely no correlation between the scenarios while a value of one indicate exact same surfaces.

Scenario	A	B	C	D	E	Z
A	1	0.99974	0.99981	0.99975	0.99981	0.99907
B	0.99974	1	0.99951	0.99999	0.9995	0.9988
C	0.99981	0.99951	1	0.99954	1	0.9993
D	0.99975	0.99999	0.99954	1	0.99954	0.99883
E	0.99981	0.9995	1	0.99954	1	0.9993
Z	0.99907	0.9988	0.9993	0.99883	0.9993	1

Figure 8: Volume changes (differences) between all the DTM surfaces (scenarios A to E) from the DTM of scenario Z. Red color indicate Gain (addition), grey represent unchanged areas and blue indicate Loss.



4 Conclusions

In this study, a semi-automated methodology for creating a DTM using Z values from points delivered from a DSM is proposed and tested on a real AoI. The implementation of six point sampling scenarios was realized using as a ground truth values 148 high precision (RTK) GCP's. Thus, the effect of point sampling quantity and point distribution on the quality of the final DTMs was examined.

Comparing the total accuracy of the DSM from the UAS-SfM pipeline which is 0.645m, with the accuracy of the DTM produced in each scenario (table 3), it is clear viewed that scenarios C, E and Z are creating DTMs of higher quality. Must be noted that scenario Z has nearly as twice as accuracy in comparison to the other two scenarios. Additionally, a close look at scenarios A and B, RMSE results leads us to the following conclusion. A random sample distribution is much more likely to generate a DTM with low accuracy rather than a one with high, mainly because the random sampling. Moreover, the ground altitude variations are not considered as a parameter during the selection of the elevation points. However, the results of scenario D indicate that a random based distribution can generate a high accuracy DTM as long as a number of high-precision GCP's is included. Additionally, scenarios C and E tend to produce DTM having similar surfaces (see table 4) and maximum correlation value of 1. This leads to the conclusion that the additional 55 high precision (RTK) GCP's had no impact when used to the stratified sampling method.

Finally, as the quantity of elevation points that inserted to the DTM generation increases the greater the resulted accuracy of the final product (fig.7)

References

Adams, S. & Friedland, C. (2011) A survey of unmanned aerial vehicle (UAV) usage for imagery collection in disaster research and management. *9th International Workshop on Remote Sensing for Disaster Response*, 15-16 September 2011. California, USA.

Chen, J., Xiang, L. & Jingjing, L. (2009) The Study of Object-Oriented Classification Method of Remote Sensing Image. *First International Conference on Information Science and Engineering*, ICISE 2009, 26-28 Dec 2009. Nanjing, China. pp. 1495–1498.

Dominici, D., Alicandro, M. & Massimi, V. (2017) UAV photogrammetry in the post-earthquake scenario: case studies in L'Aquila. *Geomatics, Natural Hazards and Risk*. [Online] 8(1), 87-103, Available from: doi: 10.1080/19475705.2016.1176605 [Accessed 21th Jan 2008].

ENVI image analysis software. (2008) *ENVI Feature Extraction Module User's Guide*. [Online] Available from: http://www.harrisgeospatial.com/portals/0/pdfs/envi/feature_extraction_module.pdf [Accessed 20th February 2018].

Heywood, I., Cornelius, S., & Carver, S. (2006). *An introduction to geographical information systems*. Harlow, Pearson Prentice Hall.

Kavroudakis, D., et al. (2018) Efficiency and Effectiveness Approaches in Spatial Data Collection of Vrisa after Lesvos Earthquake. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Gi4DM, 18–21 March 2018, Istanbul, Turkey. 42(3W4), 277–281.

Li, Z., Zhu, C. & Gold, C. (2005). *Digital Terrain Modeling: Principles and Methodology*. Washington, D.C., CRC Press.

Soulakellis, N., Tataris G., Papadopoulou E., Chatzistamatis S., Vasilakos C., Roussou O., **Papakonstantinou A.**, Kavroudakis D., Kalloniatis C. (2019) Synergistic Exploitation of Geoinformation Methods for Post-Earthquake 3D Mapping and Damage Assessment, In: Orhan, A., Chandra, M., Sunar, F., Tanzi, T.J. (eds.) *Intelligent Systems for Crisis Management: Gi4DM 2018 (Lecture Notes in Geoinformation and Cartography)*. Springer.

Soulakellis, N., et al. (2018) Synergistic Exploitation of Geoinformation Methods for Post-Earthquake 3D Mapping of Vrisa Traditional Settlement, Lesvos Island, Greece. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Gi4DM, 18–21 March 2018, Istanbul, Turkey. 42(3W4), 491-498.

Watson, D.F. & Philip, G.M. (1985) A Refinement of Inverse Distance Weighted Interpolation. *Geo-Processing*, 2(4), 315–327.