PROCESSING & WORKING WITH LIDAR DATA AND PLÉIADES IMAGES:

AN ASSISTANCE MANUAL FOR ARCHAEOLOGISTS

Authors: Jana Ameye, Mario Hernandez, Tim Van de Voorde

belspo

geography
department
Ghent University

GHENT UNIVERSITY

TABLE OF CONTENTS

1   INTRODUCTION

The aim of the LIMAMAL project (lidar data assessment for Mesoamerican archaeological landscapes), which is funded by the Belgian Science Policy Office (BELSPO) under STEREO III contract number SR/02/212, was to examine the application potential of Pléiades imagery and light detection and ranging (lidar) technologies for prospection and visualization of Mesoamerican archaeology. Nowadays, airborne lidar and satellite remote sensing are exploited for their capabilities to produce maps and visualizations based on large area high-resolution digital elevation models (DEM) [1] and as such they could be a potentially valuable resource for archaeologists.

Archaeological remote sensing has a long history in the Maya region with many pioneering efforts advocating the wide applicability of new technologies [2]. It has been a tool for archaeologists in the Maya lowlands since Alfred Vincent Kidder (1885-1963) took the first aerial photographs of the region in 1929 [3]. Viewing the published studies about the contemporary, innovative work in this region, the incorporation of remote sensing data and methods into the Maya research has met with varying results [2]. Nevertheless, each additional study refines the knowledge of both the capabilities and limitations of the available resources for the field of archaeology. In summary, geodata can be very useful for archaeologists, but it must be recognized that there are technical barriers to its use by non-experts in earth observation.

From a case study on the fusion of lidar data and Pléiades images we developed an assistance manual for the creation of visualizations by archeologists and practitioners without earth observation expertise. In a neotropical context like the Maya lowlands, lidar data offer the potential of a full-coverage survey even in densely forested areas where ground-based surveys are impractical. Given the difficulty to map the exact terrain and to identify archaeological structures, this technology with its continued advances has brought many opportunities for researchers in this region. At the same time, satellites that provide stereoscopic images with a high spatial resolution, such as Pléiades, make it possible to produce photo-realistic 3D terrain visualizations that can assist archaeologists in interpreting the landscape context. As combining both technologies can provide even more opportunities for analyzing archeological landscapes, this manual also describes the possibilities for supplementing lidar data with imagery from Pléiades..

To maximize the potential of these guidelines, we discuss the bottlenecks we encountered in this case study that impede the uptake of these two technologies. Still, the users must become familiar with the input data and the processing parameters relevant to their own research and

study site in order to achieve the best possible results The different software programs and geographic information systems (GIS) are only tools that cannot be blamed for their misuse and potentially bad results. Practice makes 'perfect'.

2 BACKGROUND KNOWLEDGE

Before starting the data fusion process of lidar and Pléiades data, it is important to have a basic understanding of the technologies we are working with. This raises the question of what the main characteristics are of these different remote sensing techniques. Therefore, the next section provides a basic explanation of the two data acquisition systems involved in the fusion.

2.1 Lidar system

Lidar is an active remote sensing technique that obtains high-accuracy three-dimensional (x,y,z) data points. The name is an acronym for *light detection and ranging,* and is used as an umbrella term for detecting and determining the distance to an object using laser light [4]. Although the technique was originally developed for military purposes (locating submarines), it became sufficiently advanced to produce highly detailed elevation models [5]. Since the 1970s and 1980s, airborne lidar systems have been used to collect topographical data. Nowadays, this leading edge laser scanning technology with its ongoing advancements is widely used in a wide range of scientific disciplines and applications: floodplain monitoring, environmental projects, geomorphology, infrastructure, archaeological research and management, etc. [6].

In order to collect airborne lidar data, a sensor can be mounted under an aircraft. This provides the ability to acquire detailed spatial data representing the form of the landscape [4, 5]. Typically, a pulse-based scanner is used in this context. This type of scanner determines distances by emitting thousands of laser pulses after which it accurately measures the elapsed time between emission and detection of the of the pulses reflected from the earth's surface [7]. Each pulse can result in multiple echoes or "*returns"* in forested areas, as they can find their way between the small open spaces of a dense canopy. The first return may thereby be received from leaves at the top of trees and vegetation, while the last return may be the actual height of the ground surface. The returns in between may then come from somewhere halfway down (e.g. branches, understory, etc.) [5,8].

Next to pulse-based scanners, there are also phase-based scanners that emit a continuous laser beam while an electronic circuit measures the phase difference between the emitted and received waveform [7]. Such scanners are, however, only rarely used for airborne laser scanning as they have a smaller measurement range.
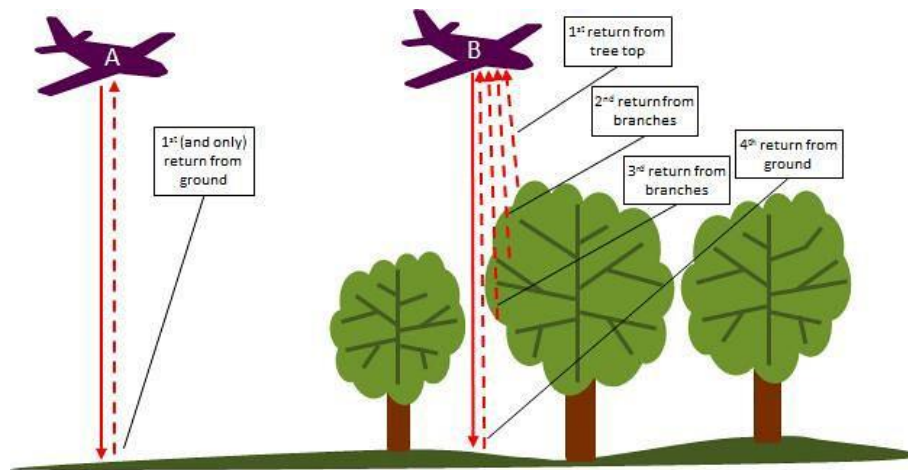
Figure 1: Principles of lidar returns in areas with and without vegetation [9].

One of the typical characteristics of airborne lidar scanning (ALS) is that it generates in a relatively short period of time a large number of terrain data in terms of resolution and scale. The number of data points easily runs into the millions, resulting in a so-called "*point cloud*" [5]. In addition to the coordinates of those points, other parameters like reflectance or intensity can be stored during the flight. Since no object information is imparted to such points, interpreting the raw data is not straightforward and this complicates working with this type of data [6]. Classification is done in post processing where the points in the cloud are assigned to a user-defined set of classes (e.g. soil, vegetation, buildings, etc.).

Nowadays, one of the key areas for using lidar technology are the wooded environments. For example, lidar data are now commonly used in archaeological research in Mesoamerica, as well in other tropical regions of the world [10]. The identification of archaeological features plays a major role in these kind of studies. Aerial lidar is therefore increasingly being used for research and applications in archaeology for investigating unknown terrain that is difficult to access, and where aerial photographs do not offer any solace [4].

2.2   Pléiades constellation

The Pléiades twins, designed by *Centre National d'Etudes Spatiales* (CNES), were developed by the French space industry with a contribution of several other European Countries. The first satellite, Pléiades 1A, was launched in December 2011, while the second, Pléiades 1B, was launched in December 2012. These very high resolution (VHR) satellites are operating in the same sun-synchronous orbit and are phased 180° from each other, which allows the satellites to revisit any point on the globe daily. As soon as an area has been collected, the images are automatically processed and quickly delivered, which makes the Pléiades constellation ideal

for monitoring emergency situations (e.g. up-to-date information from the ground, managing oil spills, risk prevention, etc.). For these applications, Pléiades can meet the most demanding requirements of their customers [11].

This satellite mission has attracted much attention due to its unique tri stereo images that can be used to create more accurate 3D models. Providing almost simultaneous images from three different viewing angles (forward, backward and nadir) for the same area, minimizes the risk of missing hidden items (Figure 2), which is ideal for dense urban and mountainous areas [11]. It allows a better retrieval of heights over terrains where the performance of classic photogrammetry is limited [1].
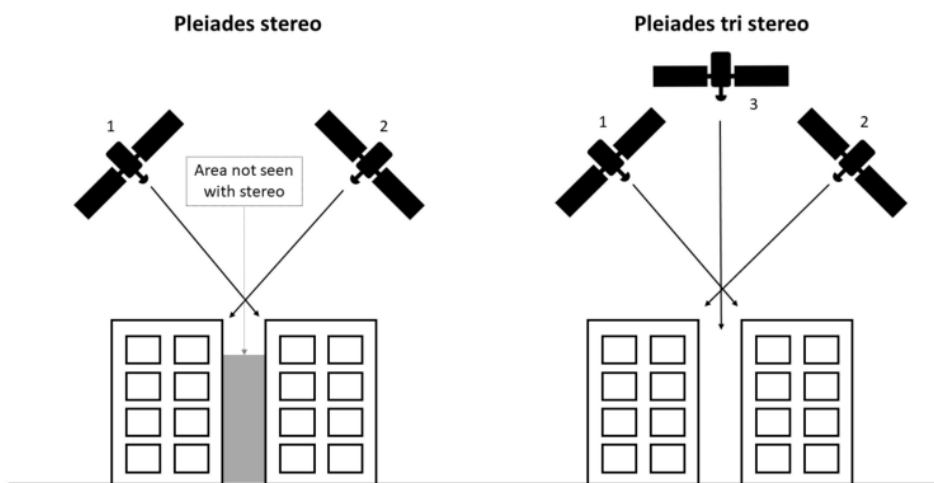

Figure 2: (Tri) Stereoscopic cover capabilities of the Pléiades constellation [1].

Regarding the sensors of the Pléiades system, the images are captured over a wide spectral band (Figure 3). It carries a four-band multispectral (MS) sensor with a maximum spatial resolution of 2 m, and a panchromatic (PAN) band with a maximum spatial resolution of 0,5 m [11]. The multi-band sensor of the MS imaging system consists of a set of detectors sensitive to the blue, green, red and near-infrared parts of the spectrum. In contrast to the PAN imaging system that consists of a single band detector sensitive to radiation within a wide spectral range covering the visible part of the spectrum, the recorded radiation of a MS system is within a relatively narrow range of wavelengths for each band. Due to lower amount of energy per unit area gathered within its narrower bandwidths, the MS images always have a lower resolution than PAN images obtained from the same satellite [12]. Other main characteristics of the Pléiades constellation are further detailed in the Pléiades Imagery User Guide [11].
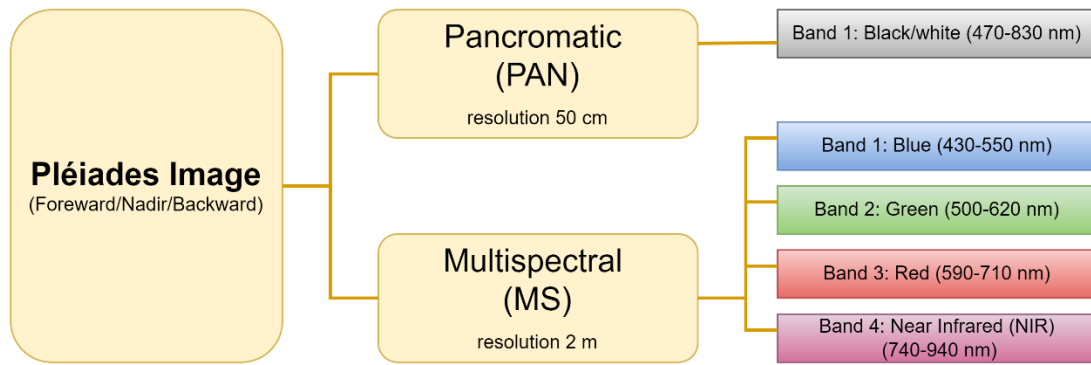
Figure 3: Pléiades channels and bands with the corresponding spectral range.

3   GENERAL INFORMATION

3.1   Software

The guidelines for LIMAMAL were developed using different software programs. These are listed below (including the version number of the program):

- ArcGIS Pro – 2.9.0

  (Paid software: https://pro.arcgis.com/en/pro-app/2.8/get-started/download-arcgis-pro.htm)

- Catalyst Professional – 2222.0.5

  (Paid software: https://catalyst.earth/products/catalyst-pro/)

- CloudCompare – 2.12

  (Free software: http://www.danielgm.net/cc/release/)

- QGIS – 2.18.28

  (Free software: https://www.qgis.org/nl/site/forusers/download.html)

- Relief Visualization toolbox (RVT) – 2.2.1

  (Free software: https://www.zrc-sazu.si/en/rvt)

- Snap – 8.0

  (Free software: https://step.esa.int/main/download/snap-download/)

3.2   Study area

For this project, an attempt was made to select an area with the presence of known archaeological Mayan vestiges. The Puuc Region of Yucatan, Mexico, a cultural subregion of the ancient Maya and a distinct physiographic zone within the Yucatan peninsula proved to

be ideally suited for this purpose. Along the road of the *Ruta Puuc* many archaeological zones can be found, both discovered and undiscovered. Since the pioneering explorations of John Lloyd Stephens and Frederick Catherwood around 1840, a (re)discovering of the ancient Maya civilization took place. They carried out a systematic investigation, recognized a regional civilization with a common architectural style from the Late Classical Period, and conveyed its wonders to a wider world [14].

Geographically, the Puuc region is an area with a low range of hills running along the northwestern part of the Yucatan peninsula. This is separated with the uplifted wedge of rolling terrain just to the south. Moreover, the soil in the valleys is deeper and more fertile than in the other areas, resulting in ideally suited areas for agriculture. In this way, food is provided not only to the inhabitants of the region, but probably also for trade within Yucatan. Eventually, a dense zone of cone karst hills lies to the South of Valle de Santa Elena which divides the landscape, but the water table was inaccessibly low. For this reason, seasonal ponds and small lakes were particularly valuable in the Puuc region and the construction of water reservoirs became necessary [15].

In addition to the better-known sites in the Puuc region (e.g. Uxmal, Labna, Kabáh and Sayil), this document focusses on the less familiar one of Kiuic. This is a well-preserved site which is under supervision of the Instituto Nacional de Antropologia e Historia (INAH) Yucatan. Kiuic is a Mayan word with two possible meanings that are linked to each other: "main square" and "market", but in ancient Mexico the market was always at the main square. Many structures have eventually been found on the visualizations we made of that site.

Figure 4: The entrance gate of Kaxil Kiuic [Mario Hernandez, UGent].

3.3   The data

Due to the need to use an existing lidar dataset, the possibilities were limited for studying a very specific area. With the availability of free, public data provided by NASA's G-LiHT Data Center Webmap (https://glihtdata.gsfc.nasa.gov/), the request for the data capture of Pléiades images was tailored accordingly. This means that a specific request for Pléiades data is not free, but there are already many VHR satellite images online (i.e. https://www.intelligence-airbusds.com/imagery/sample-imagery/, https://www.pleiades4belgium.be/, etc.).

3.3.1   Lidar

The free lidar data used in these guidelines was collected by environmental scientists over extensive areas of Mexico in April and May 2013, using NASA-Goddard's lidar Hyperspectral & Thermal Imager (G-LiHT) system that seeks to promote full and open data sharing [16]. The primary objective of this mission under the direction of Dr. Bruce Cook was to refine the measurements of aboveground carbon stocks in the Mexican forest [17]. Care was taken to release accurate, well-calibrated data in a timely manner and to distribute it in a common, easy-to-use file format such as LAS, the typical format for lidar point clouds [18]. This file format is also a standardization of the American Society for Photogrammetry and Remote Sensing (ASPRS) in which lidar point clouds are divided into different classifications (e.g. ground, building, water, etc.) [19]. From the more than 400 tiles available through the G-LiHT Data Center Web map, the data from Kuiic1_Apr2013 was downloaded. According to the metadata, the data were collected with a Riegl Laser Measurement system. The sensor was flown on a Piper Cherokee N4118R aircraft at a nominal altitude of 335 m above ground level and a nominal velocity of 110-150 kt [16]. Some other interesting values are shown in Table 1 as technical background. The data themselves are obtained as a TAR and GZ file format, respectively the predecessor of a zip file and a compressed collection of one or more files. Each point in the raw LAS file has been given the following parameters in addition to its coordinates: point source id, scan angle rank, edge of flight line, number of returns, return number, gps-time, intensity and classification.
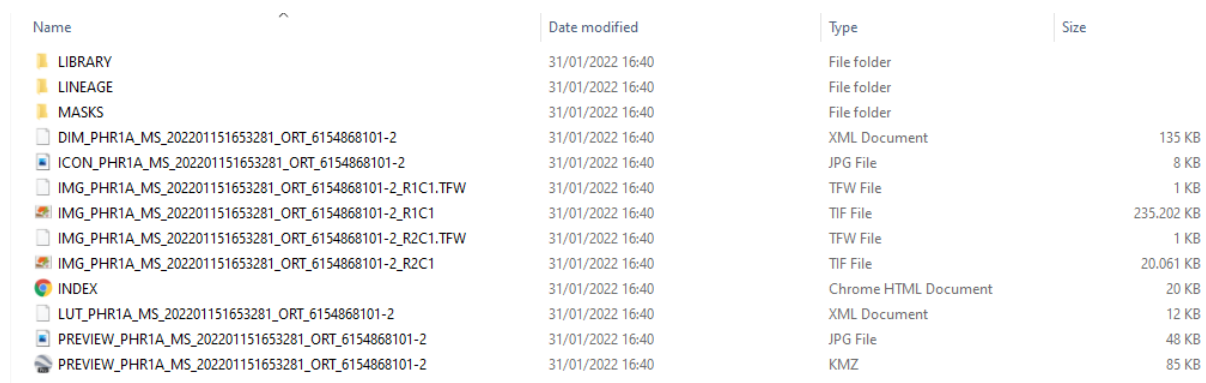
Table 1: Metadata Kuiic1_Apr2013 [16].

Lidar equipment	VQ-480 (Riegl)
Laser wavelength	1550 nm
Pulse width	3 ns
Beam divergence	0.3 mrad
Laser pulse repetition frequency	300 kHz
Nominal distance between points in a scan line	0.24 m
Nominal distance between scan lines	0.56 m

3.3.2 Pléiades

The purchased Pléiades images for this project were ordered through the Belgian Science Policy Office (BELSPO) from Airbus, the supplier that then makes the images available via an FTP connection. Using a script, these files are automatically downloaded to BELSPO's database, after which they could be obtained via a OneDrive using the appropriate credentials. Altogether, the tri stereo images and the orthophoto take up a total of 2.46 GB. The images date from January 15, 2022 and contain no clouds over the area of interest.

Some of the information of the Pléiades images can be determined from the file name. For example, 'MS' indicates that it is a multispectral product, while 'P' stands for a panchromatic one. The raster products are supplied by the vendor as a set of files in a product folder. When browsing to the different product folders, there are typically a variety of files included (Figure 5). As can be seen, there are several files and some additional folders. The XML Document includes the metadata and the instructions needed to display the product in a GIS software. The software reads the XML document to determine which files to use and how to use them exactly. The XML file thus contains a considerable amount of information. The TIF File is the actual image file that was collected by the sensor. It may not have any georeferencing information associated directly with it, or only a very coarse one, causing it to appear in the wrong place after loading it into a GIS application. Finally, the HTML Document is a simplified and easier-to-read metadata file that opens in a Web browser [20].

Name	Date modified	Type	Size
LIBRARY	31/01/2022 16:40	File folder	
LINEAGE	31/01/2022 16:40	File folder	
MASKS	31/01/2022 16:40	File folder	
DIM_PHR1A_MS_202201151653281_ORT_6154868101-2	31/01/2022 16:40	XML Document	135 KB
ICON_PHR1A_MS_202201151653281_ORT_6154868101-2	31/01/2022 16:40	JPG File	8 KB
IMG_PHR1A_MS_202201151653281_ORT_6154868101-2_R1C1.TFW	31/01/2022 16:40	TFW File	1 KB
IMG_PHR1A_MS_202201151653281_ORT_6154868101-2_R1C1	31/01/2022 16:40	TIF File	235.202 KB
IMG_PHR1A_MS_202201151653281_ORT_6154868101-2_R2C1.TFW	31/01/2022 16:40	TFW File	1 KB
IMG_PHR1A_MS_202201151653281_ORT_6154868101-2_R2C1	31/01/2022 16:40	TIF File	20.061 KB
INDEX	31/01/2022 16:40	Chrome HTML Document	20 KB
LUT_PHR1A_MS_202201151653281_ORT_6154868101-2	31/01/2022 16:40	XML Document	12 KB
PREVIEW_PHR1A_MS_202201151653281_ORT_6154868101-2	31/01/2022 16:40	JPG File	48 KB
PREVIEW_PHR1A_MS_202201151653281_ORT_6154868101-2	31/01/2022 16:40	KMZ	85 KB

Figure 5: Content of a Pléiades multispectral product folder.

3.4   Important remarks

There are some important issues that need to be taken into account when reading and using this assistance manual on lidar data and Pléiades stereoscopic image processing (data fusion) for archaeological purposes:

- The lidar data and the Pléiades images we used in our case study are not from the same date. There is a gap of nine years that should be taken into consideration while processing and analyzing the data.

- In normal circumstances, data are linked to a project and not the other way around. In this case, the area focused on data that was already available, or in other words the study area was adjusted accordingly.

- Using public data may come with some constraints such as the limited availability, a specific reference system, etc. Nevertheless, it is free and it may help for a more targeted data collection as it can give a first impression of the region. In the areas of interest, more data can then be acquired with the desired resolution and bands.

- Whereas the lidar data that we used in this case study was available for free from NASA's Goddard's Lidar, Hyperspectral and Thermal Imager (G-LiHT - https://gliht.gsfc.nasa.gov), it is by no means the best and most recently available lidar dataset for mapping archeological structures in the area. A more recent dataset with considerably less noise was collected by the National Center for Airborne Lidar Mapping (NCALM) in 2017 using the newer Teledyne Optech Titan multispectral lidar sensor [21,22]. This dataset is still under embargo as it is currently being analyzed by W. Ringle and colleagues who conduct research in the Puuc region [22]. There is yet another lidar dataset available that was collected by the Alianza México REDD+. This dataset could not be found online or by using the contacts at INAH.

- In contrast to the G-LiHT lidar data, the high resolution stereoscopic satellite images we used were purchased from Airbus Industries (a commercial provider) with funds provided by the Belgian Science Policy Office. Although some high resolution images may be available for free, in general new acquisitions and in most cases also archive data have to be purchased. This means that the workflows presented in this document that rely on stereoscopic images (such as the visualization of pansharpened, multispectral images on a lidar-derived digital surface model) will require some degree of funding.

4 PROCESSING

In this chapter, the different processing steps are explained. The workflow that we applied in this case study consists of three parts (figure 6): methods for lidar data, methods for Pléiades imagery and methods for the data fusion. How to obtain the various sub-products is carefully explained step by step along with the corresponding actions that need to be performed in the different software packages. As a reader, these packages give you the opportunity to choose between a number of common alternatives, both commercial and free. This makes it possible to generate results for other fields of study as well, using this manual. We will first discuss how to derive a digital elevation model from lidar and tri stereoscopic data and then how to create visualizations and orthophotos based on those elevation models. In addition, we will discuss some other products that can be derived from Pléiades data such as a pansharpening, a true color composite and a false color composite. We will also discuss how to distinguish lidar ground points from other points.
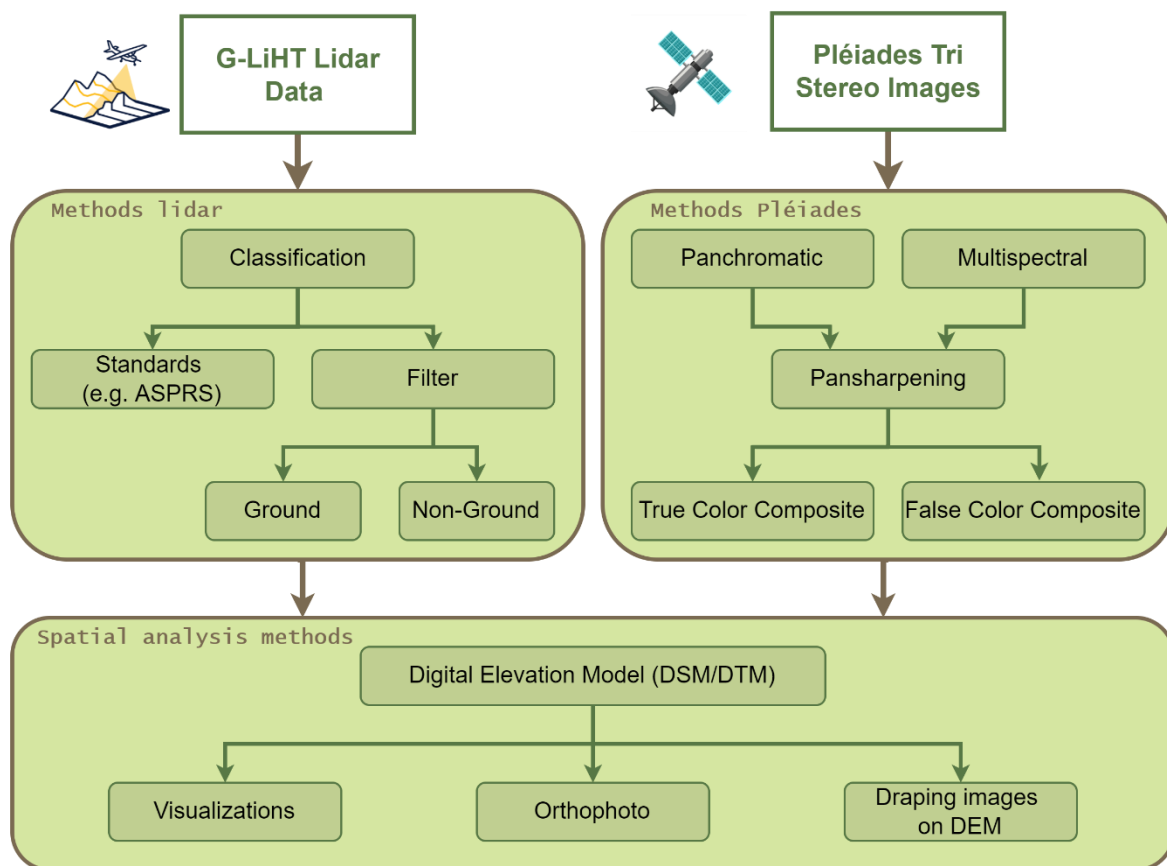

Figure 6: Processing in the data fusion case study: lidar data + Pléiades imagery.

4.1 Creating digital elevation models

It is possible to produce a high-quality digital elevation model (DEM) from both lidar data and Pléiades images. However, two different kinds of DEMs can be reconstructed: the digital surface model (DSM) and the digital terrain model (DTM). The first reproduces a regular and continuous representation of the elevation surface of the landscape and all objects (e.g., buildings, trees, etc.) located in the investigated area, while the second reproduces only the elevation surface of the bare ground [23]. Both models are normally characterized by high planimetric and altimetric accuracy, and are often the input for further visualizations such as orthophotos, color composites, hillshading, etc. This is also be demonstrated further in this case study.

To obtain these digital elevation models, the workflows below rely on different software. First, *Catalyst Professional* is used for the creation of a DSM from Pléiades images. This is a PCI Geomatics brand that is designed for earth observation image processing and delivers scalable business solutions using the power of spatial data and analytics [24]. A free trial with full functionality is available for seven days. Second, the 3D point cloud processing software from *CloudCompare* is used to process the lidar data from the G-LiHT data center. This is an open source project and thus freely available. It provides a set of basic tools for manual processing and rendering of point clouds supplemented by some different advanced processing algorithms [25]. *CloudCompare* can generate a DTM as well as a DSM. When the digital elevation models are obtained from CloudCompare, it is advisable to provide the correct coordinate reference system (crs) in a GIS program, such as QGIS (free, open source software), as the grid will not be located correctly in comparison with other map layers. In this case study, the crs of the original lidar dataset was used, namely WGS84 / UTM zone 16N (EPSG 32616). Finally, creating DEMs is also possible with the third software package we used, namely *ArcGIS Pro*, which is a well-known, powerful, licensed desktop GIS application developed by Esri. It supports data visualization, advanced analysis and data maintenance in both 2D, 3D and 4D [26]. A workflow to create the difference-maps used in section 5 was also added below.

WORKFLOW PLÉIADES DSM EXTRACTION IN CATALYST PROFESSIONAL [27]

Initial Project Setup

1. Open the Geomatica *OrthoEngine* application from Catalyst Professional
2. Click *File > New*
3. Give the project a *Filename*, *Name* and *Description*
   - Select *Optical Satellite Modelling* as the Match Modelling Method
   - Select *Rational Function* (Pléiades) under Options
   - Click *OK*
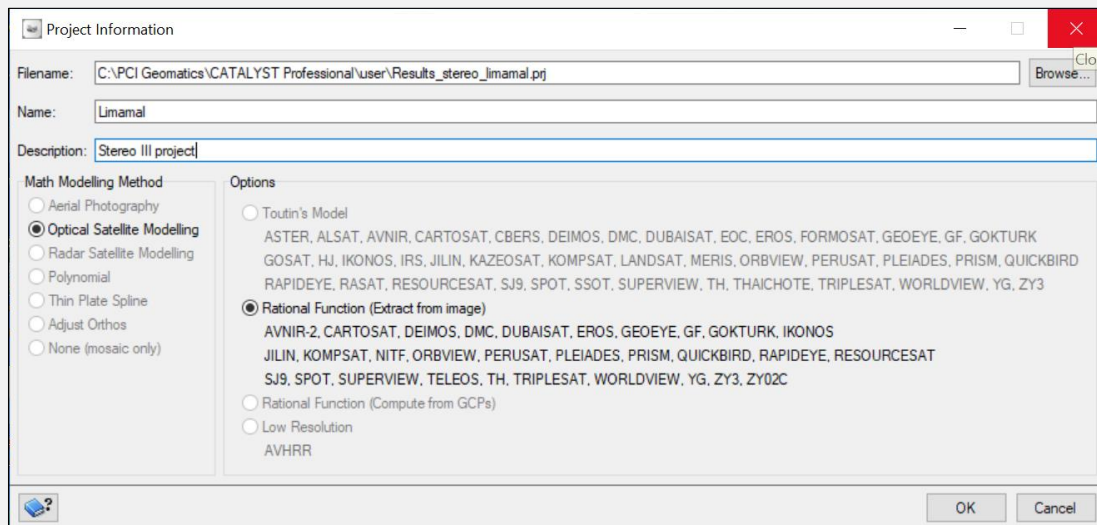
The Project Information interface is used in order to create a new OrthoEngine project (GUI from Catalyst Professional - PCI software).

Data Input

1. Select *Data Input* in the Processing step
2. Click *Open a new or existing image*
   - Click *Add Image*, navigate to the location of the data and select the xml-file in the image folder of the raw stereo image
   - Select *Yes* when asked if you want to import data file to PIX file to optimize processing
   - Select a *file name* and *location* for the output pix file
   - Select *Yes* when asked if you want to create overviews now
   - Repeat the previous steps to add all images to the project[1]
   - Click *Close*

---

[1] This includes all stereo images. For the tri stereo Pléiades images, these are the three different viewing angles for the panchromatic band as these images have the highest resolution.
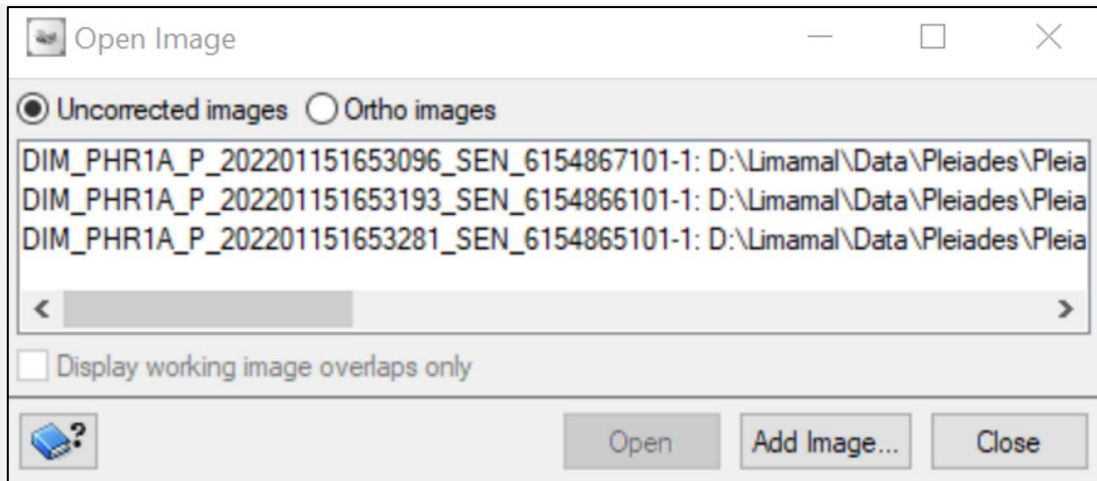
The Open Image interface is used to open the three different viewing angels from the panchromatic, tri stereo Pléiades images (GUI from Catalyst Professional - PCI software).

Collect Tie Points

1. Select *GCP/TP Collection* in the Processing step

2. Click *Automatically Collect Tie Points*

3. Leave the parameters as defaults

4. Click *Collect Tie Points* (TP)[2]

5. Close the Automatic Tie Point Collection window

6. Click *Display Project Overview* to see the location of the used tie points

   ▪ Click with the right mouse on the images > *Show* > *TPs*
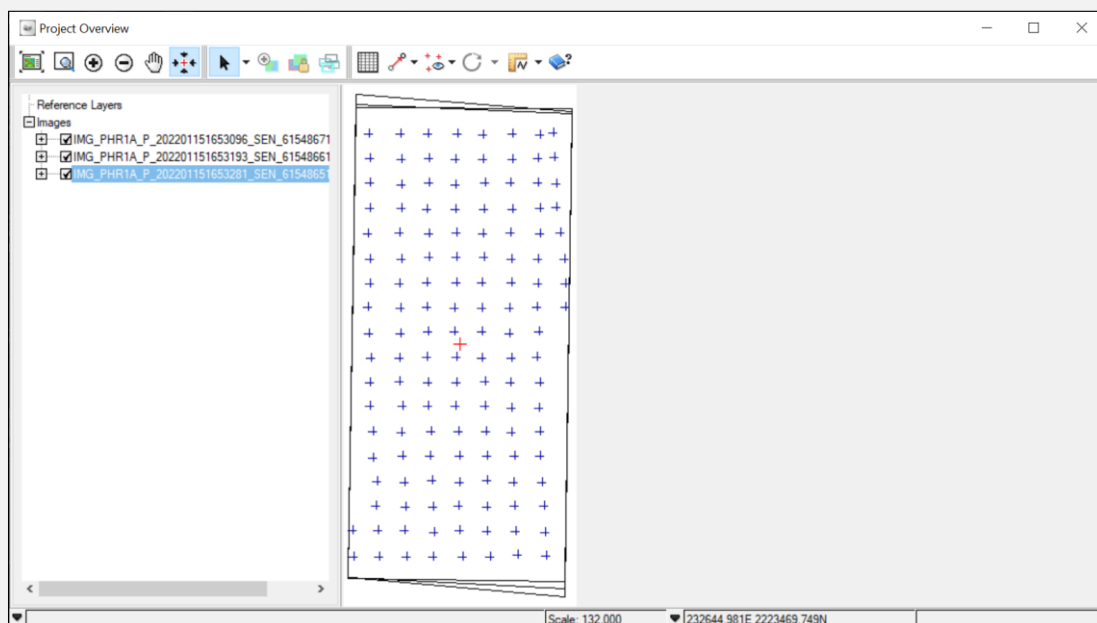
The Project Overview interface is used to visualize the tie points on the different viewing angles (GUI from Catalyst Professional - PCI software).

---

[2] Tie points are common objects or locations within the overlap areas between the tri stereo images [27].

Refine Tie Points

1. Select *Model Calculations* in the Processing step
2. Click *Compute model*
3. Click *OK* to the pop-up message that states that the Block Adjustment has completed
4. Select *GCP/TP Collection* in the Processing step
5. Click *Residual report*
   - The collected TPs are displayed in the window
   - If you have any points with very high residual errors, you can remove them from the model using the *Delete Point* button. This is not the case in this example.
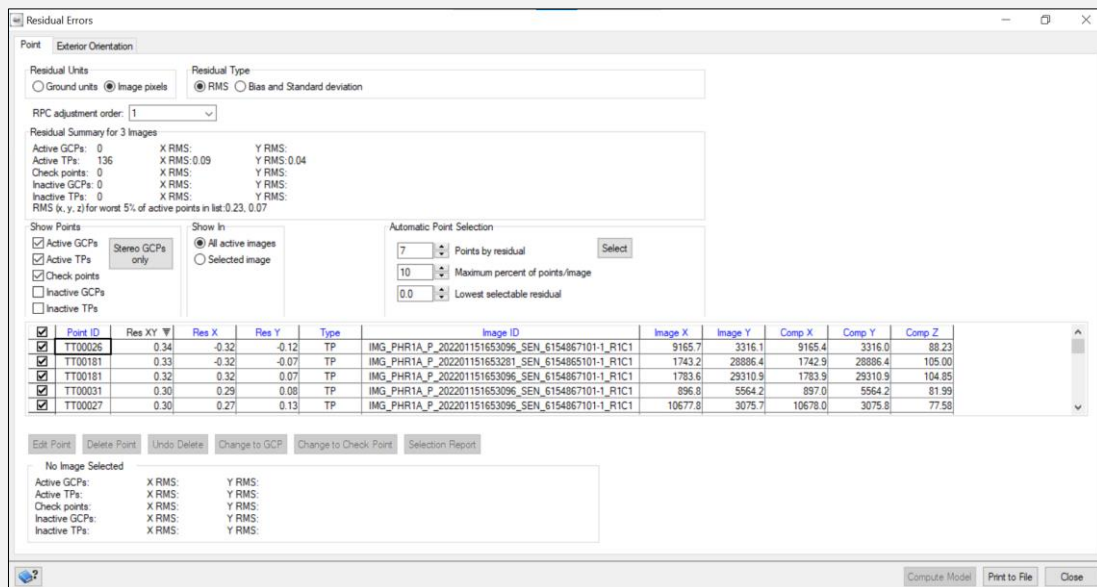
The Residual Errors interface is used to remove low-quality 3D tie points. These points have high residual errors (GUI from Catalyst Professional - PCI software).

DSM Generation

1. Select *DEM From Stereo* in the Processing step
2. Click *Create Epipolar Image*[3]
   - Select left and right image (3 times for tri stereo Pléiades images)
   - With both images selected, click *Add Epipolar Pairs To Table*
   - Repeat the previous two steps for the three different possibilities with tri stereo data
   - Click *Generate Pairs*
   - Click *OK* to the pop-up massage that states that the epipolar generation is completed successfully and close the Generate Epipolar Images panel

---

[3] Epipolar images are stereo pairs that are reprojected to have a common orientation for the left and the right image. The matching features between the images appear along a common x-axis [27].

3. Click *Extract DEM automatically*
   - Select the three epipolar pairs by checking the Select box
   - Under DEM Extraction Options
     i. Select *SGM* (Semi-global matching)[4]
     ii. Check the Output DEM vertical datum
     iii. Check mark *Epipolar tracking*
   - Under Geocoded DEM
     i. Select *Create geocoded DEM*
     ii. Select an *output file name* and *location*
     iii. Set the *X* and *Y* resolution (e.g. 1 m)
   - Click *extract DEM*
   - Click *OK* to the pop-up message that states that the DEM is extracted successfully
   - Close the Automatic DEM Extraction panel

Save Geotiff

1. Open the saved file in the Geomatica *Focus* application from Catalyst Professional
2. Right mouse on the DSM > *Save As*
   - Select an *output file name* and *location*
   - Select the *TIFF 6.0 Format*
   - Click *Save*
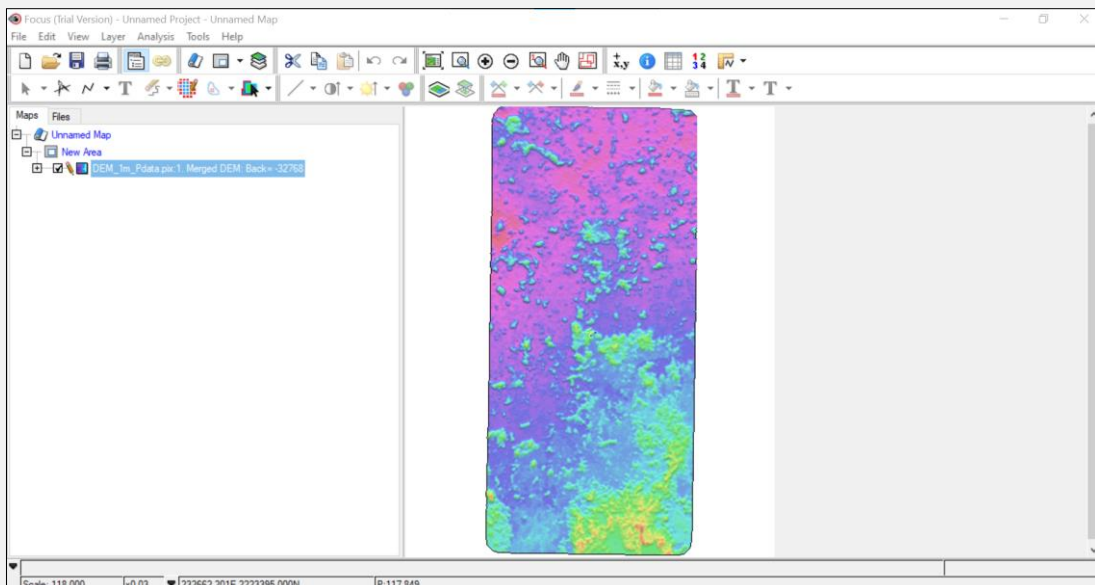
The Focus interface is used to create a TIFF file from the generated digital elevation model (GUI from Catalyst Professional - PCI software).

---

[4] Semi-global matching is a computer vision algorithm for the estimation of a dense disparity map from a rectified stereo image pair [27].

WORKFLOW LIDAR DSM EXTRACTION IN CLOUDCOMPARE

Project Setup

1. Open the *CloudCompare* application
2. Click *Open*, navigate to the location of the data and select the appropriate LAS-file in the image folder
3. Select *Apply all* in the Open LAS file panel
4. Select *Yes to All* in the Global shift/scale panel
5. Select the individual LAS files (if any) to *Merge* under the Edit tab
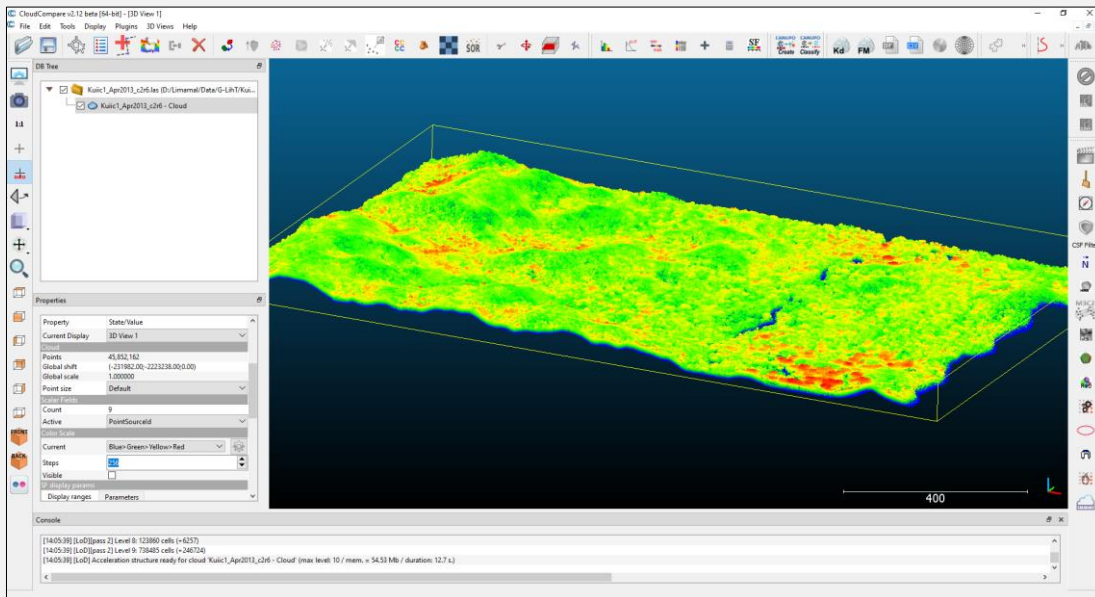
Point clouds from Kiuic opened in a new project and visualized according to the PointSourceId of the source data (GUI from CloudCompare).

DSM Generation

1. Select the appropriate file which must be processed into a grid
2. Select *Rasterize* (convert a cloud to 2D raster) under Tools/Projection
    - Under Grid:
        i. Choose the grid step: 1 m
        ii. Active layer: Height grid values
    - Under Projection:
        i. Direction: Z
3. Click *Update grid*
4. Select *Raster > Export heights > OK*
    - Select an *output file name* and *location*
    - Click *Save*

<h1 style="text-align:center">WORKFLOW LIDAR DTM EXTRACTION IN CLOUDCOMPARE</h1>

Project Setup

See explanation Project Setup under 'workflow lidar DSM extraction in *CloudCompare*'.

Further processing requires the bare ground points for the extraction of the DTM. There are two possibilities to obtain these ground points. Either a filter (CSF) is applied to obtain the ground points, or an already existing classification included in the LAS file is used. Depending on the objectives of the own project, you select the most appropriate method.

Ground Classification Filter (<u>optional</u>)

1. Select the appropriate point cloud for the area of interest
2. Select the Cloth Simulation Filter (CSF)[5] under the Plugins tab
   a. Under General parameter setting:
      i. Select the appropriate scene (e.g. Steep slope)
   b. Under Advanced parameter setting:
      i. Select the appropriate Cloth resolution[6], Max iterations[7] and Classification threshold[8] (e.g. 0.5, 500, 0.5)
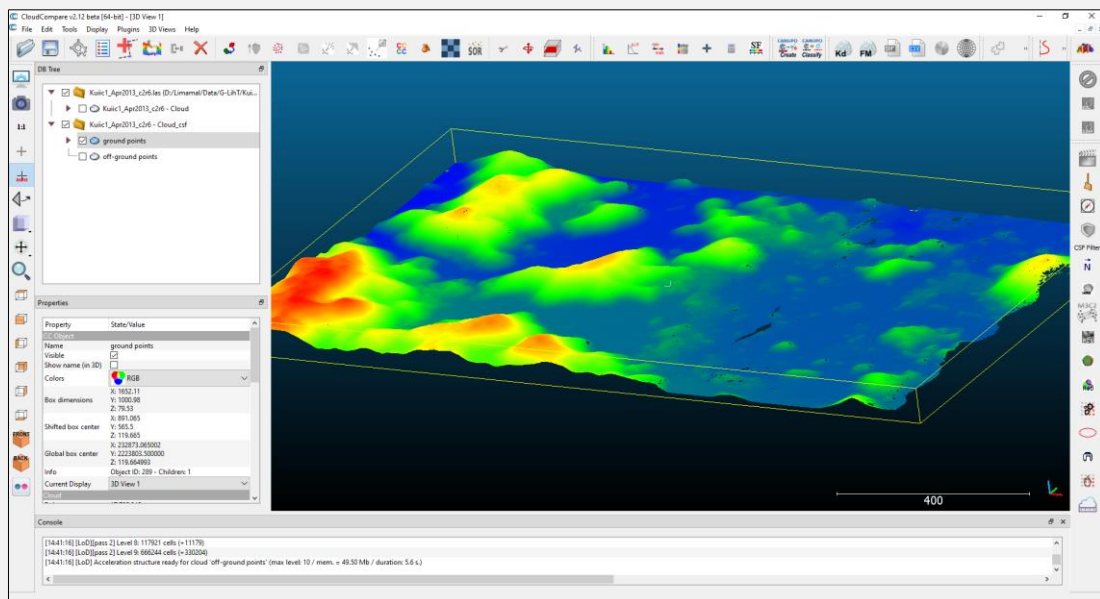
The CloudCompare interface shows the ground points obtained after running the Cloth Simulation Filter (GUI from CloudCompare).

---

[5] With CSF, the original point cloud is turned upside down, and then a simulated cloth falls on the inverted surface from above, dividing the point clouds into ground and non-ground parts [28].

[6] Cloth resolution refers to the grid size of cloth which is used to cover the terrain [28].

[7] Max iterations refers to the maximum iteration times of terrain simulations [28].

[8] Classification threshold refers to a threshold to classify the point cloud into ground and non-ground parts based on the distances between points and the simulated terrain [28].

Extraction of Ground Points in the Existing Classification (<u>optional</u>)

1. Select the appropriate point cloud for the area of interest
2. Activate *Classification under Scalar Fields* in the Properties box
3. Select *Filter by Value*
   - Range: 1.5 – 2.5
   - Click *Export*
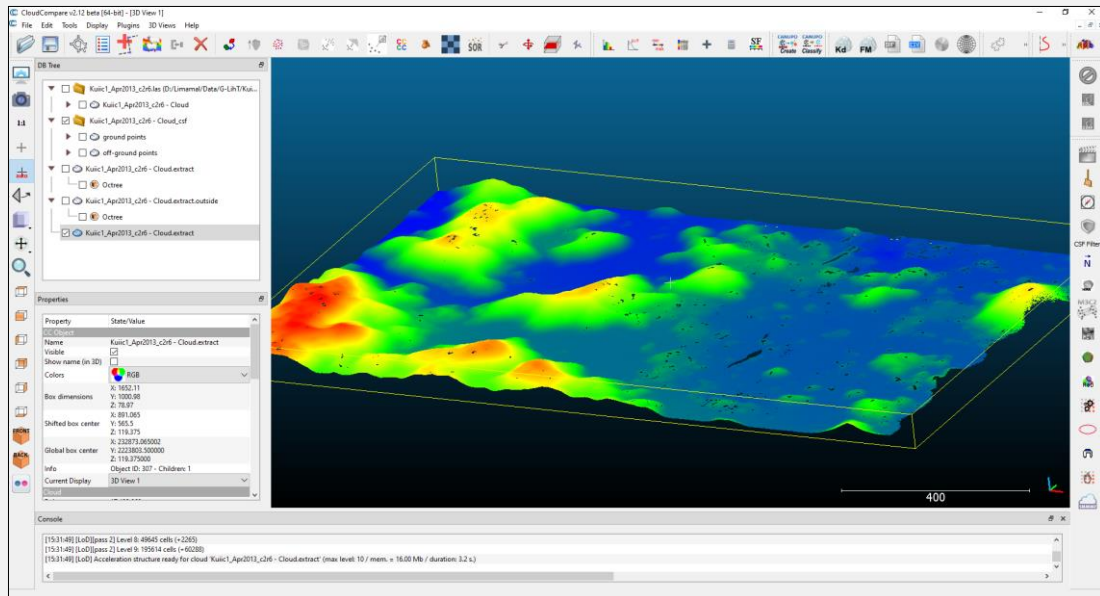
The CloudCompare interface shows the ground points obtained after running Filter by Value to extract the existing classification (GUI from CloudCompare).

DTM Generation

See explanation DSM Generation under 'workflow lidar DSM extraction in *CloudCompare*'.

When selecting the appropriate point cloud, the files with the ground points are selected.

WORKFLOW ADDING A CRS TO A RASTER IN QGIS

Creation of a Raster Layer with a Coordinate Reference System

1. Open the QGIS software on your desktop
2. Click *Layer > Add Layer > Add Raster Layer*
   - Select the *Input raster*
   - Click *Open*
3. Select the appropriate system in the Coordinate Reference System Selector panel
   - Search in the Filter for the correct EPSG (e.g. 32616)
   - Select the correct Coordinate Reference System
   - Click *OK*

4. Right-click the layer in the Layer Panel > *Save as*
   - Browse to a *location* and give a *name* to the output raster
   - Click *Save*
   - Click *OK*
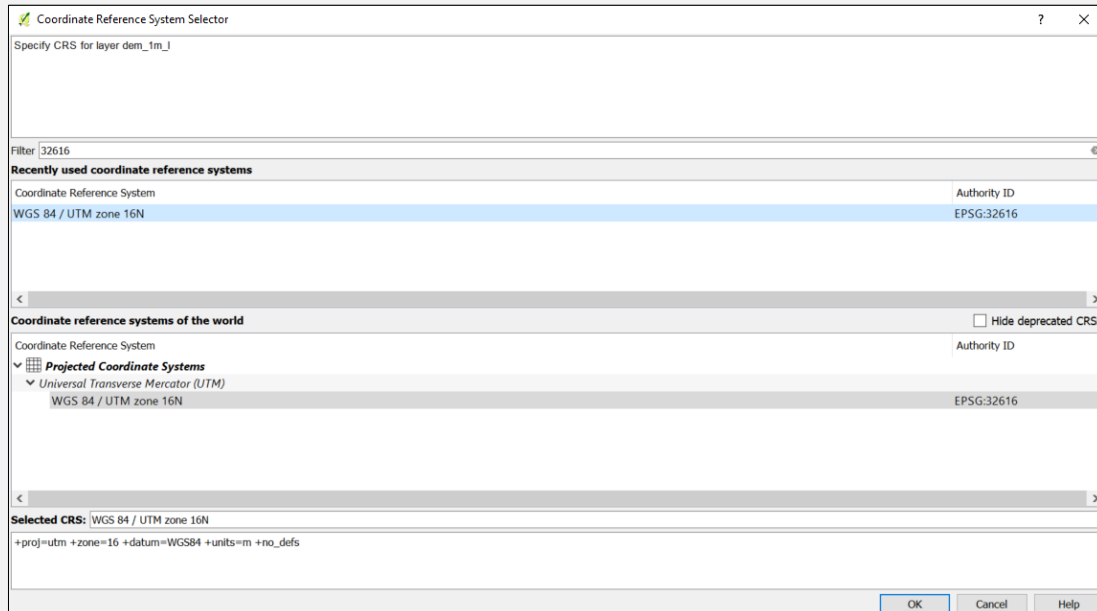
The Coordinate Reference System Selector interface is used to select the appropriate reference system, in this case WGS 84 / UTM zone 16N (GUI from QGIS).

**WORKFLOW LIDAR DEM EXTRACTION IN ARCGIS PRO.**

Project Setup

1. Create a *New Project* (Local Scene) in ArcGIS Pro
2. Give the project a *Name* and a *Location*
3. Click *OK*

Create a LAS Dataset

1. Select the *Tools button* in the Analysis tab
2. Search *Create LAS Dataset*[9]
   - Select the *Input Files*
   - Give the Output LAS Dataset a *Name* and *Location*
   - Select the appropriate *Coordinate System*
   - Click *Run*

---

[9] The LAS dataset is a stand-alone file that resides in a folder and references lidar data in the LAS format with optional surface constraint features that define surface characteristics [29].
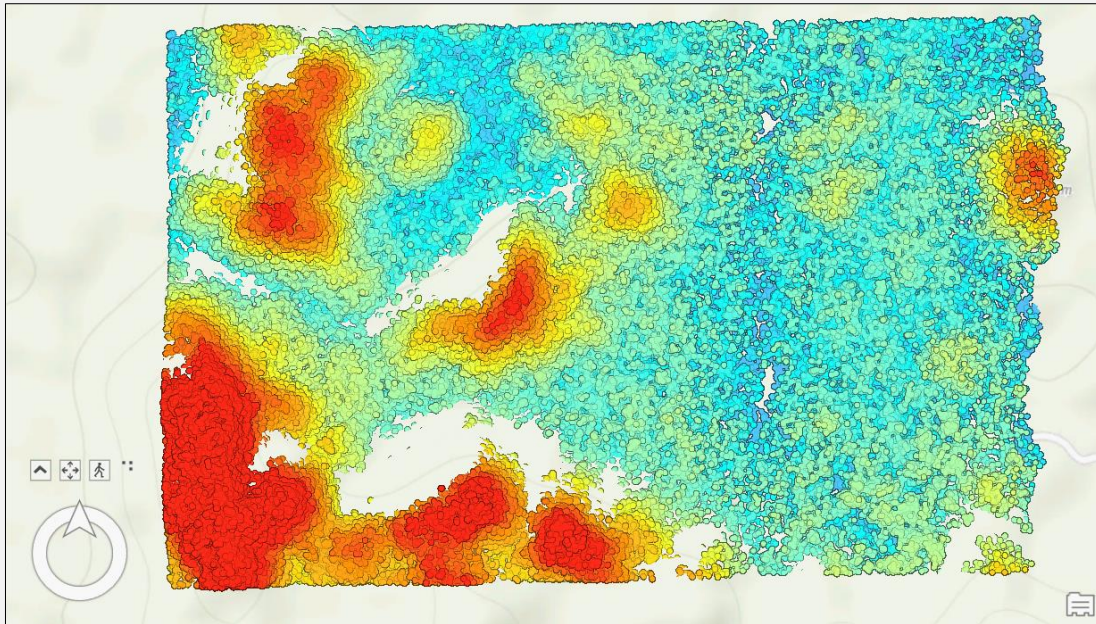
The map view interface from ArcGIS Pro that displays the created LAS dataset of Kiuic (GUI from ArcGIS Pro).

DEM Generation

With the DEM generation of ArcGIS Pro, both a digital surface and a digital terrain model can be generated. Since we are working with a LAS dataset, it is possible to use the already existing classification for the selection of the ground points. As in CloudCompare, filters can also be applied here.

1. Right-click on the created dataset in the Contents Pane and select *Properties*
2. Go to the *LAS Filter* section and select the desired option:
    - **DTM**: Only the Ground Points
    - **DSM**: All Points
    - Click *OK*
3. Click on *Tools* (Geoprocessing) in the Analysis tab
4. Search for *LAS Dataset To Raster* (Conversion Tools)
    - Select the created LAS dataset for Input LAS Dataset
    - Set a *Name* and *Location* for the Output Raster
    - Set the *Value Field* to Elevation
    - Choose the appropriate *Sampling Value*[10] (e.g. 1)
    - Click *Run*

---

[10] The Sampling Value is a parameter to define the resolution of the output raster [30].
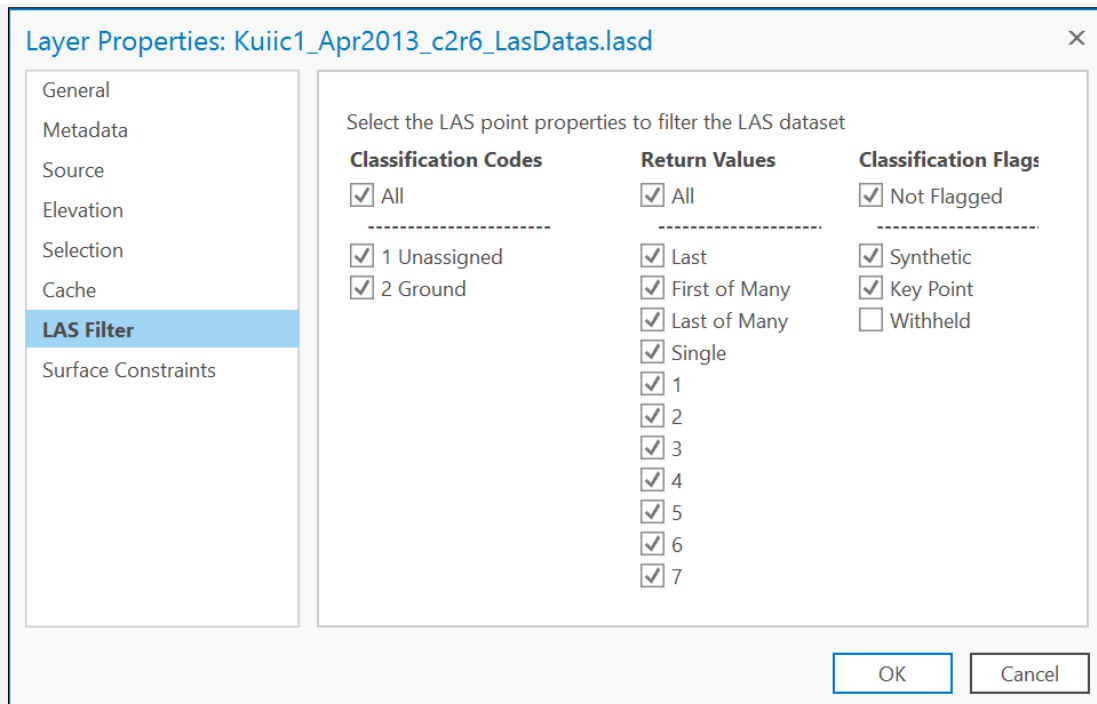
The Layer Properties panel from ArcGIS Pro to filter the appropriate points in the LAS dataset to generate a digital elevation model (GUI from ArcGIS Pro).

WORKFLOW DIFFERENCE DEMs IN QGIS

Difference between DEMs

For further processing, the differences between the obtained digital elevation models were examined. These images can be found in the results section, but here is a brief explanation of how to achieve this. For example, the differences between the obtained DTMs are examined as there are differences depending on the algorithm used.

1. Open the QGIS software on your desktop
2. Select *Raster Calculator[11]* in the Raster Tab
    - Browse to a *location* and give a *name* to the Output layer
    - Select the desired *Output format*
    - Use the Raster bands and the Operators to create the appropriate expression
    - Click *OK*

---

[11] The Raster Calculator tool allows you to create an execute a map algebra expression that will output a raster [31].
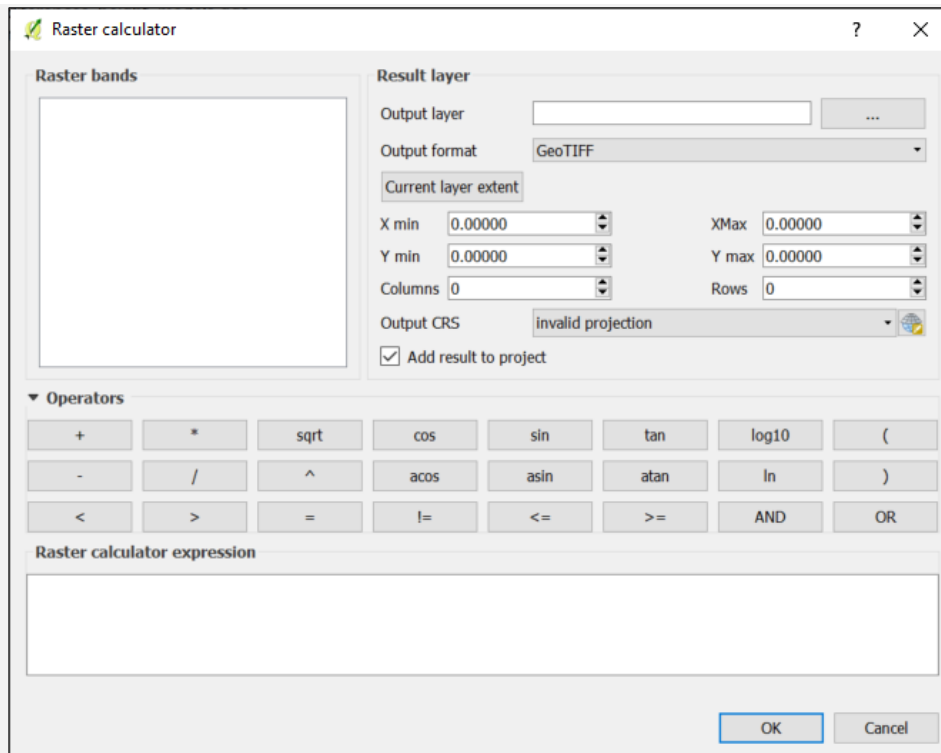
The Raster calculator panel to execute a map algebra expression (GUI of QGIS).

Transparency

1. Right-click on the layer and select *Properties*
   - Go to the *Transparency* section
   - Move the bar for the desired percentage of transparency
   - Click *OK*

Basic Map Export

1. Click on the *Composer Manager* in the Project tab
   - Click *Add* and give the new composer a *Name*
   - Click *OK*
   - Click *Show*
2. The page layout appears:
   - Click *Add New Map* and draw a box, the map will appear in it
   - Click *Add New Legend* and draw a box, the legend will appear in it
   - Click *Add New Scalebar* and draw a box, the scalebar will appear in it
   - Click *Add New Label* and draw a box, a title can be entered
   - Click on *Export as Image* in the Composer tab
   - Browse to a *location* and give a *name* to the Image
   - Click *Save*

4.2   Visualizing a digital terrain model

The creation of terrain visualizations is very useful valuable for characterizing relief features in the landscape. A digital terrain model can be visualized by applying several techniques . Choosing the most appropriate visualization method will depend on the study area and the purpose of the study. Each analytical model to visualize the data has its benefits and drawbacks for the identification of those features, making a good choice extremely important. Some models are better at highlighting features in low relief topography while others are better in defining steeply sloping earthworks. For that reason, it is also interesting to create multiple models to ensure that all archaeological features can be identified [5].

This guide provides some more information about three frequently used types of visualization, but also shows the possibility of deriving many others using the indicated software (e.g. Relief Visualization Toolbox). The methods described below are hillshading, slope and sky view factor.

Hillshading	Hillshading or relief shading is the most widely used technique for displaying relief. It provides the most natural visual impression of all the techniques. Shades of gray improve the perception of the morphology. Typically, the surface is illuminated by a direct, simulated light that is constant for the entire dataset. The relief shadows thus depend on the position of the light source, which is determined by the azimuth and elevation. By varying these parameters, different lighting effects can be achieved. In this process, areas perpendicular to the light beam are the most illuminated, while areas that cannot see the light directly are in shadow. As a result, this visualization is mainly useful in mountainous areas [32,33].
Slope	This visualization method calculates the steepness (degrees or percentage) of each cell of the grid surface to visualize it. The lower the slope value, the flatter the terrain. Conversely, a steeper terrain has a higher slope value. The inverted gray scale benefits the display of morphology, where it is very suitable for detecting local level differences. This also makes the visualization quite easy to interpret [33,34].
Sky view factor	This technique uses uniforme, diffuse illumination. This is very useful in archaeology because it improves the recognition of small-scale features. In addition, this method can solve the directional problems of hillshading. It measures the portion of the sky visible from a given point, but only if the elevation data are not manipulated with a vertical exaggeration. Thus, relief illumination is correlated with a portion of the visible sky that is limited by the relief horizon. As a result, a point on a ridge is brighter than a point at the bottom of a steep valley. The explanation is that a larger portion of the sky can be seen on the ridge [32,33].

To obtain the visualizations, the workflows below use two different software tools: *ArcGIS Pro*, which was already introduced in Section 4.1, and the *Relief Visualization Toolbox* (RVT). The latter is free software developed by the Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU) to help scientists visualize grid elevation models such as the previously created DEMs. It offers a limited selection of techniques that have proven to be effective for the identification of small-scale, archaeological features with default settings assuming that the project is working with high-resolution DEMs derived from lidar scan missions [35].

WORKFLOW FOR DTM VISUALIZATIONS IN ARCGIS PRO

Project Setup

1. Create a *New Project* (Map) in ArcGIS Pro
2. Give the project a *Name* and a *Location*
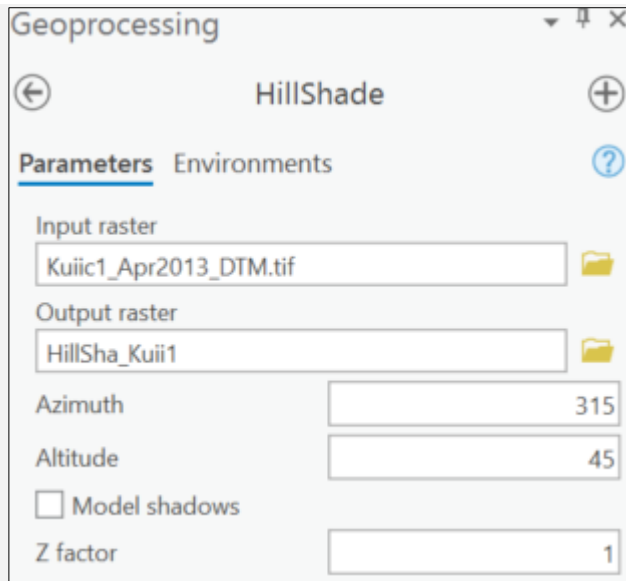3. Click *OK*
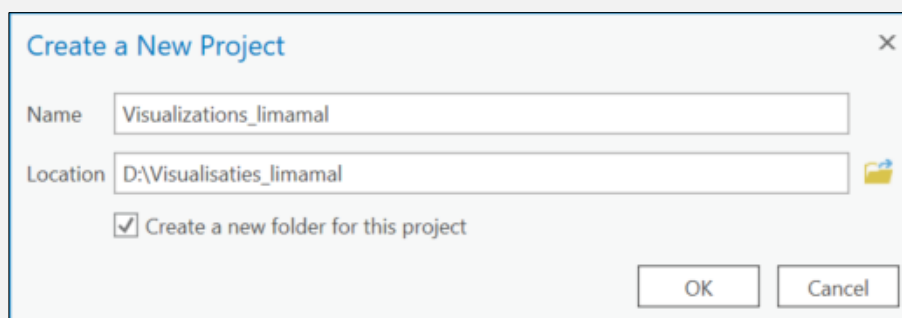
Panel to create a new project in ArcGIS Pro (GUI from ArcGIS Pro).

Hillshade Model

1. Click on *Tools* (Geoprocessing) in the Analysis tab
2. Search for *HillShade* (3D Analyst Tools)
   - Select the *Input raster*
   - Select a *location* and give a name to the *output raster*
   - Choose values for the different parameters
     - Azimuth[12] (e.g. 315)
     - Altitude[13] (e.g. 45)
   - Click *Run*

---

[12] The azimuth is the angular direction of the sun, measured from north in clockwise degrees from 0 to 360. An azimuth of 90 degrees is east [36].
[13] The altitude is the slope or angle of the illumination source above the horizon [36].

The HillShade panel from the Geoprocessing toolbox to create a hillshade (GUI from ArcGIS Pro).

Slope Model

1. Click on *Tools* (Geoprocessing) in the Analysis tab
2. Search for *Slope* (3D Analyst Tools)
   - Select the *Input raster*
   - Select a *location* and give a name to the *output raster*
   - Choose the appropriate parameters (e.g. degree)
   - Click *Run*
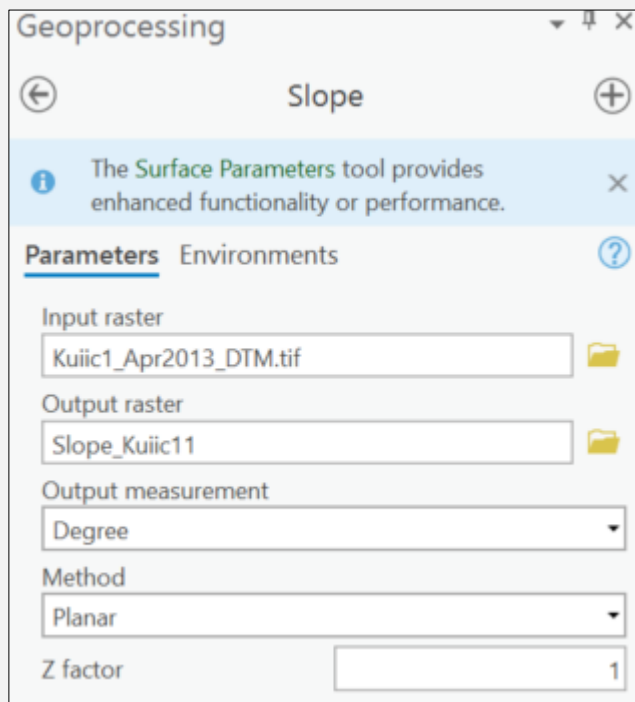
The Slope panel from the Geoprocessing toolbox to create a slope visualization (GUI from ArcGIS Pro).

WORKFLOW FOR DTM VISUALIZATIONS IN RVT

Creation Visualization Models

1. Open RVT and select *Click to continue*
2. Click *Add File(s)* in the RVT panel
3. Tick the checkboxes of the desired visualizations and click *Start*
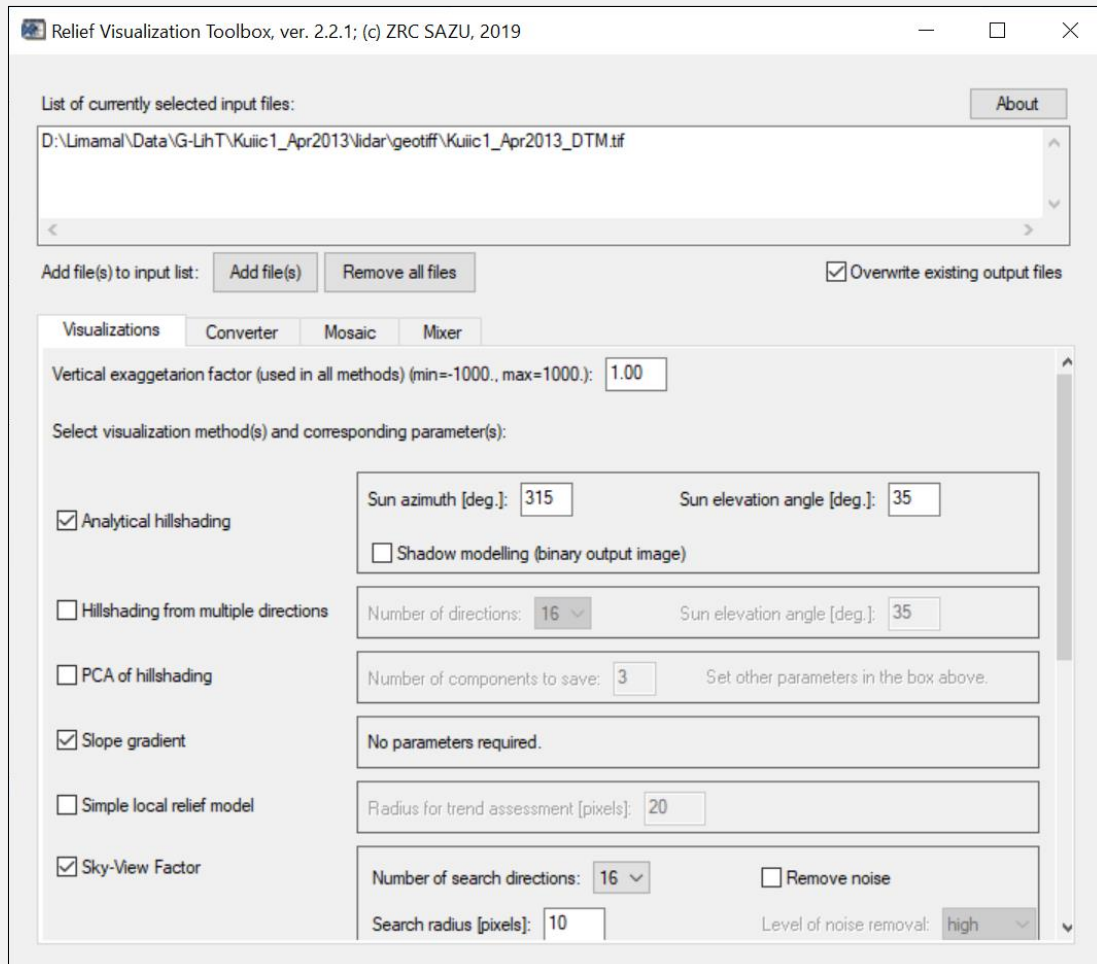
The RVT panel is used for making different kind of visualizations for the case study. There are more possibilities than just hillshading, slope and sky view factor (GUI from the Relief Visualization Toolbox).

4.3 Creating an orthophoto

Orthorectification of the Pléiades images removes the systematic geometric distortion induced by the central projection, the relief and the not-verticality of the recording axis. The ortho product is a georeferenced image that has been corrected for acquisition and terrain off-nadir effects (parallax) [11]. As a result, this image is very valuable as basic information for surveying plots as it is more accurate metrically. The process of orthorectification is carried out in the workflows below using *ArcGIS Pro* and *Catalyst Professional*.

WORKFLOW FOR CREATION OF ORTHO IMAGES IN ARCGIS PRO.

Project Setup

1. Create a *New Project* (Map) in ArcGIS Pro
2. Give the project a *Name* and a *Location*
3. Click *OK*

Generate Ortho Image

1. Click on *Tools* (Geoprocessing) in the Analysis tab
2. Search for *Create Ortho Corrected Raster Dataset*[14] (Data Management Tools)
   - Select the *Input raster* (one of the tri-stereo images)
   - Select a *location* and give a name to the *output raster*
   - Choose the appropriate *Orthorectification Type* (e.g. DEM)
   - Select the *DEM Raster*
   - Click *Run*
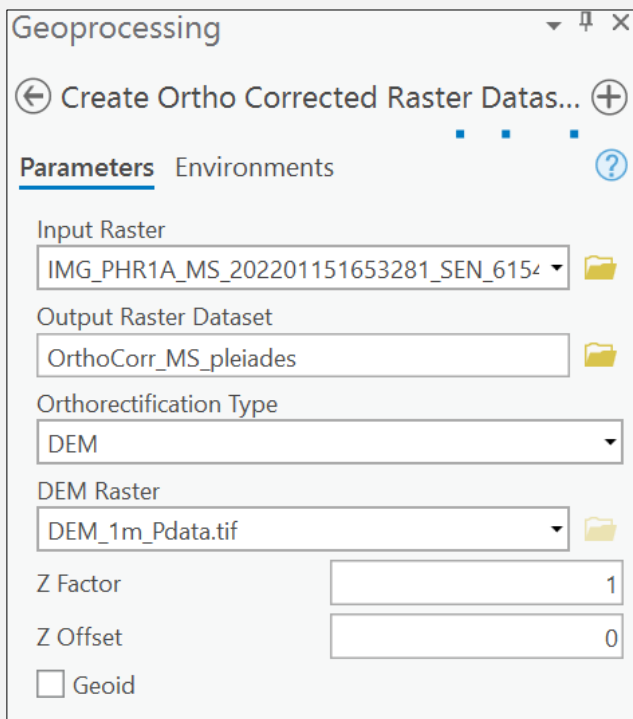
The Create Ortho Corrected Raster Dataset panel from the Geoprocessing toolbox to create a slope visualization (GUI from ArcGIS Pro).

---

[14] The Create Ortho Corrected Raster Dataset is a geoprocessing tool that creates an orthocorrected raster dataset using the rational polynomial coefficients associated with satellite imagery [37].

WORKFLOW FOR CREATION OF ORTHO IMAGES IN CATALYST PROFESSIONAL [38]

Project Startup

First follow all the steps under 'workflow Pléiades DSM extraction in *Catalyst Professional*'.

Generate Ortho Images

1. Select *Ortho Generation* in the Processing step
2. Click *Schedule Ortho Generation*
   - Select the Available images (the image of which an orthophoto should be made) in the left panel and add them to the images to process panel
   - Browse to the output location for your ortho images and select a DEM file
   - Click *Generate Orthos*
   - When all Processing Completed, click *Close*
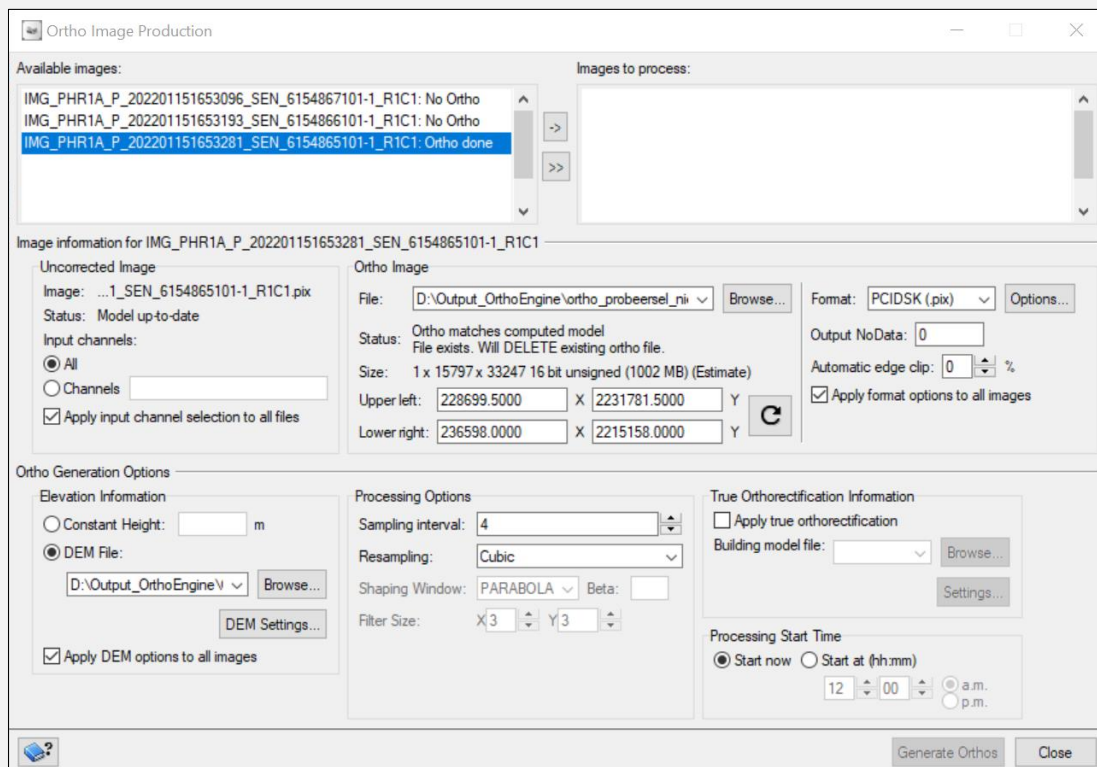
The Ortho Image Production interface to produce orthophotos (GUI from Catalyst Professional - PCI software).

4.4 Creating a Pansharpened image

Pansharpening is a process of fusing the visual colored information of the multispectral data with the spatial details provided by the panchromatic data to create a single high spectral and spatial resolution image. The advantage of this better spatial resolution facilitates the identification of objects [12]. Typically, for the multispectral data, three of the four low-resolution bands are used as the main inputs in the process [11]. This makes it possible to generate a pansharpening even for false color images (Section 4.5). Finally, different results are possible by working with different pansharpening methods. The workflows below use the *Catalyst Professional* and *ArcGIS Pro* software.

WORKFLOW FOR PANSHARPENING IN ARCGIS PRO
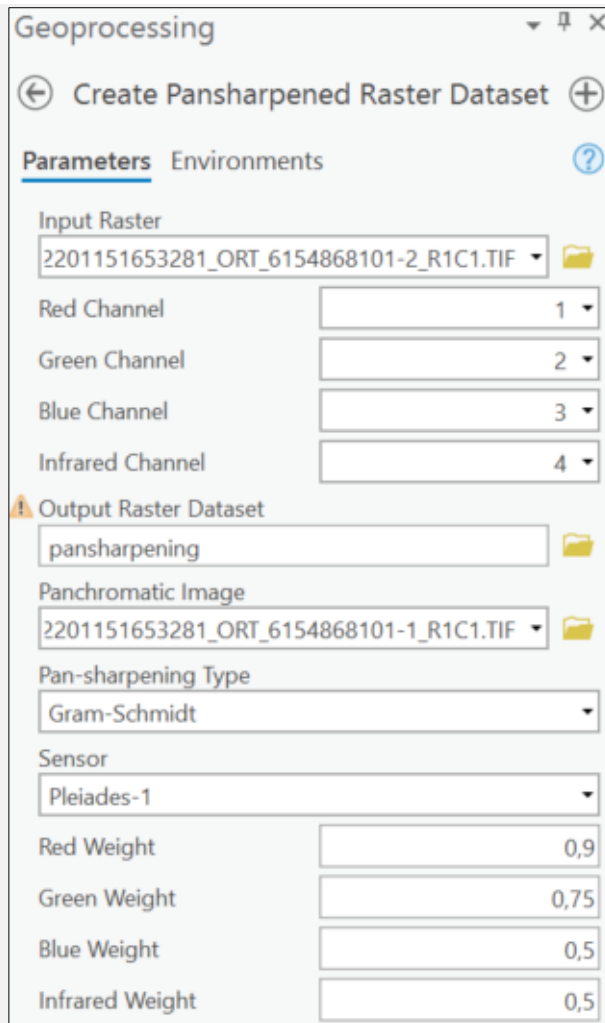
Project Setup

1. Create a *New Project* (Map) in ArcGIS Pro
2. Give the project a *Name* and a *Location*
3. Click *OK*

Pansharpening

1. Click on *Tools* (Geoprocessing) in the Analysis tab
2. Search for *Create Pansharpened Raster Dataset* (Data Management Tools)
    - Select the *Input raster* with the correct Channels
    - Select a *location* and give a *name* to the Output Raster Dataset
    - Select the *Panchromatic Image*
    - Choose the appropriate *Pansharpening Type*[15] (e.g. Gram-Schmidt)
    - Select the correct *Sensor*
    - Click *Run*

---

[15] Many different pansharpening types are possible [39]:
- The Brovey algorithm based on spectral modelling for data fusion.
- The Esri algorithm based on spectral modelling for data fusion.
- The Gram-Schmidt algorithm to sharpen multispectral data.
- IHS that uses intensity, hue and saturation color space for data fusion.
- Mean that uses the average value between the green, red and blue values and the panchromatic value.

The Create Pansharpened Raster Dataset panel from the Geoprocessing toolbox to create a pansharpening (GUI from ArcGIS Pro).

WORKFLOW FOR CREATING A PANSHARPENED IMAGE IN CATALYST PROFESSIONAL [40]

Project Setup

See explanation Initial Project Setup under 'workflow Pléiades DSM extraction in *Catalyst Professional*'.

Pansharpening

1. Click on *Tools > Merge/Pansharp Multispectral Image*
    - Select the *Multispectral images* in the dataset
    - Select the *Panchromatic image*
    - Select the *output filename* and *location* for the output PIX file
    - Click *Pansharp*
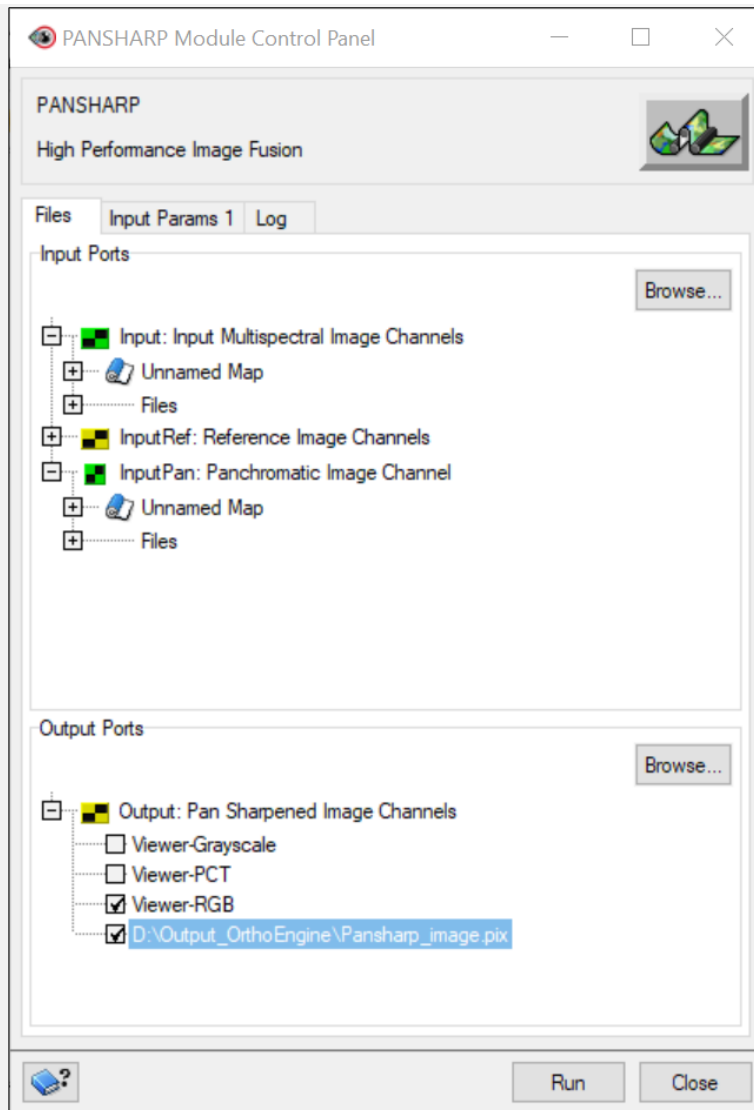
The PANSHARP Module Control Panel to create a pansharpening (GUI from Catalyst Professional - PCI software).

4.5   False color composite

In addition to true color images (RGB), it is also possible to generate false color composites. These images are a representation of a multispectral image produced by using different bands than the visible red, green and blue as the components of an image display. In other words, these images give the ability to visualize wavelengths that the human eye cannot see [41]. In the case of Pléiades, this is only the near infrared band (740-940 nm). Using image processing software, it is easy to create such a false color image. In the procedures below, the images are rendered using *ArcGIS* Pro and *SNAP*. This last software is a Sentinel Application Platform, a common architecture for all Sentinel toolboxes. It is ideal for Earth Observation processing and analysis [42], even for other satellites than Sentinel.

WORKFLOW FOR THE CREATION OF FALSE COLOR COMPOSITES IN ARCGIS PRO

Project Setup

1. Create a *New Project* (Map) in ArcGIS Pro
2. Give the project a *Name* and a *Location*
3. Click *OK*

Creation False Color Composite

1. Right-click on Map in Contents and Select *Add Data*
   - Navigate to the *location* of the data
   - Select the *input file*
   - Click *OK*
2. Right-click on the Red/Green/Blue channel and choose the appropriate band

WORKFLOW FOR THE CREATION OF FALSE COLOR COMPOSITES IN SNAP

Creation False Color Composite

1. Open the SNAP application
2. Select *File > Open Product…*
   - Navigate to the *location* of the data
   - Select the *input file*
   - Click *Open*
3. Right-click on the opened file and select *Open RGB Image Window*
   - Select the appropriate channels for Red, Green and Blue
   - Click *OK*
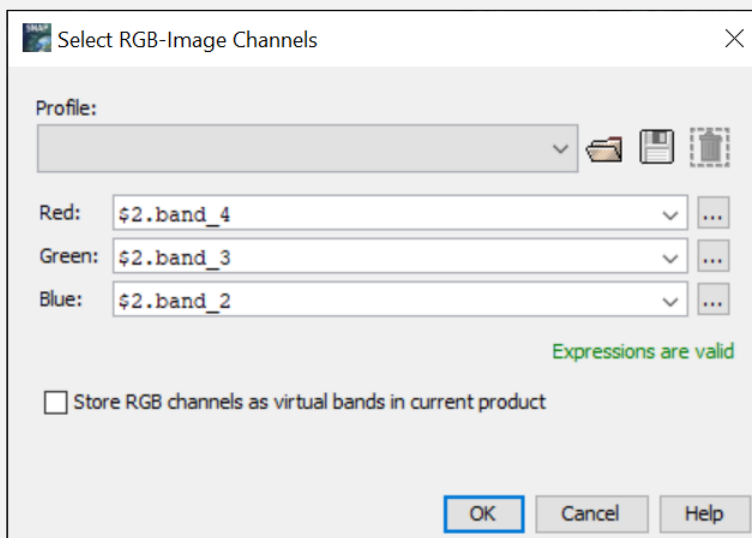
The Select RGB-Image Channels panel to select the bands at a given color (GUI from SNAP).

4.6 Draping a basemap over lidar data

To create a more realistic landscape, it is possible to drape any basemap over a digital elevation model created from a lidar LAS dataset in ArcGIS Pro. This gives plenty of possibilities to create interesting visualizations, which you can see below.
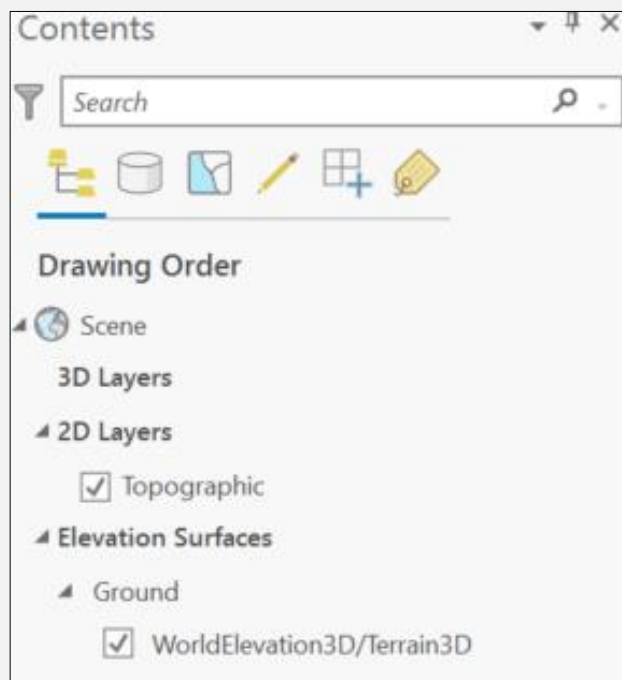
WORKFLOW DRAPING AN IMAGE OVER LAS DATA IN ARCGIS PRO [43]

Project Setup

1. Create a *New Project* (Global Scene) in ArcGIS Pro
2. Give the project a *Name* and a *Location*
3. Click *OK*

Draping Image over DEM

1. Right-click *Elevation Surfaces* in the Contents pane and choose *Add Elevation Surface*

The contents pane shows the different scenes of which Elevation Surfaces is one (GUI from ArcGIS Pro).

2. Right-click on the newly created surface and choose *Add Elevation Source*
3. Browse to the created DEM surface from the previous step
4. Select *Add Data > Data* in the Map tab to open the desired Image

5. Right-click on the image in the 2D Layers and select *Properties*
    - Go to the Elevation section
    - Features are.. "*on custom elevation surface*"
    - Choose the appropriate *Custom Surface*[16]
    - Choose the appropriate *Cartographic offset*[17] (e.g. 25)
    - Click OK
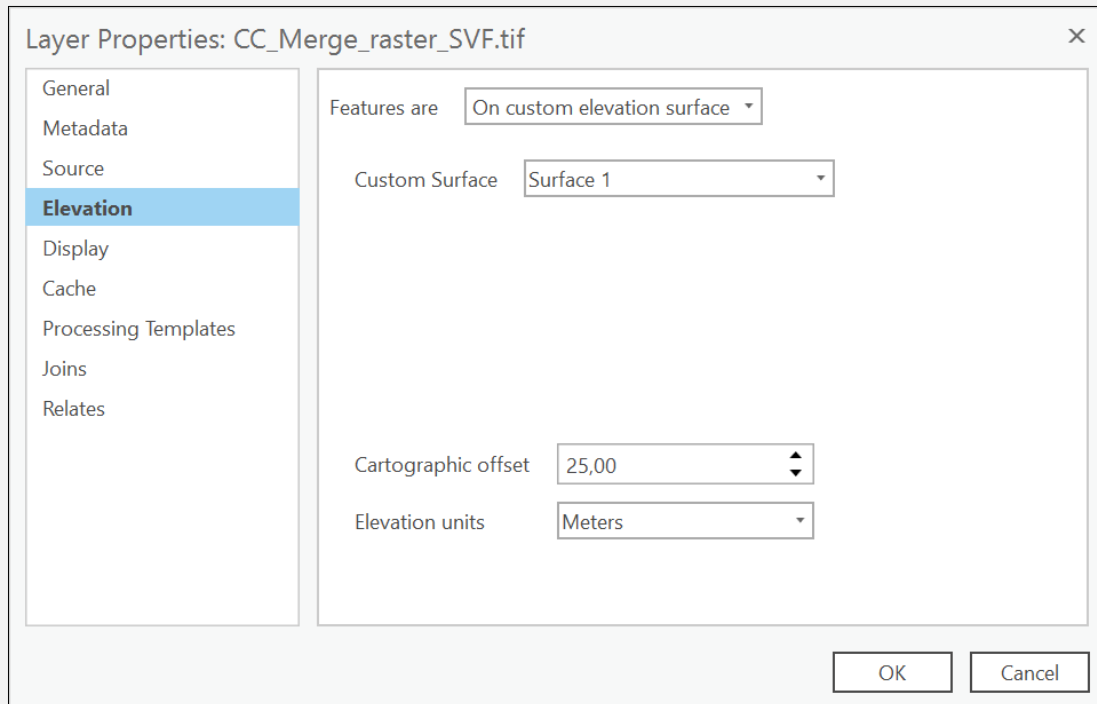
The Layer Properties Panel for draping a basemap on an elevation model (GUI from ArcGIS Pro).

---

[16] A custom surface is a surface that defines the depth of the terrain [44].
[17] A cartographic offset vertically adjusts the z-value of the entire layer [44].

5 RESULTS AND DISCUSSION

5.1 Creation of elevation models

From both the lidar data and the Pléiades images, we created several elevation models. Depending on the desired result, the appropriate processing method and optimal parameters should be selected. It was not our intention to provide an exhaustive overview of all potentially useful software with these guidelines, only to demonstrate some data fusion methods for archaeological purposes. Should the methods used in this manual do not meet the expectations of your project, we recommend you to experiment with other software, which you will certainly find online.
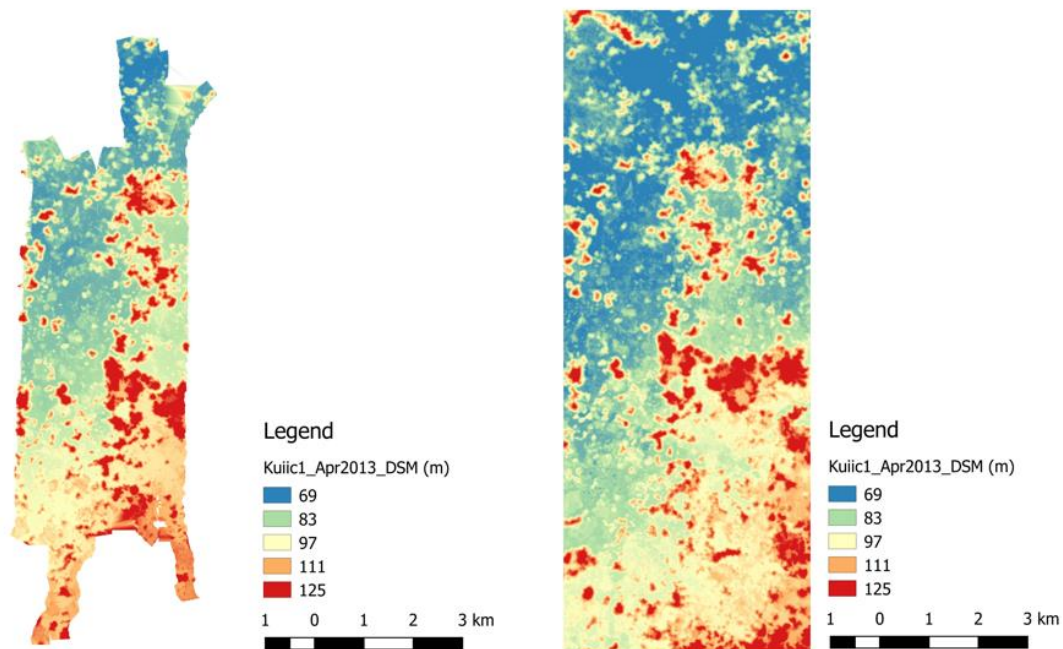

Figure 7: DSM of the area of interest: G-LiHT lidar data (a, left) and Pléiades images (b, right).

Regarding the creation of DSM's, the processing of the Pléiades images turned out to be quite tedious because finding suitable processing software proved difficult. In most cases, the software had to be paid for (e.g. *Erdas Imagine*, *Catalyst professional*) or, if it was free, it was not so easy to install or use. For more advanced GIS users, we recommend checking out the NASA *Ames Stereo Pipeline* (ASP) software. ASP is free and open source automated software designed for processing (tri) stereo imagery captured from satellites around the Earth and other planets to produce cartographic products and 3D models [45]. The disadvantage of ASP is that the installation is not provided for Windows, but only for Linux and Mac. Many

users are excluded because of this, unless an additional working environment is created on Windows. Processing the lidar data in *CloudCompare* or *ArcGIS Pro*, on the other hand, is somewhat simpler. However, to process large areas in *CloudCompare*, a better processing time can be achieved when using the command prompt instead of the general user interface that has to visualize everything. In any case, a powerful computer is required to work with point clouds. The final results of the DSM derived from the Pléiades imagery and the lidar data can be found in Figure 7. On this figure you can also see the overlap between the satellite images and the lidar data.
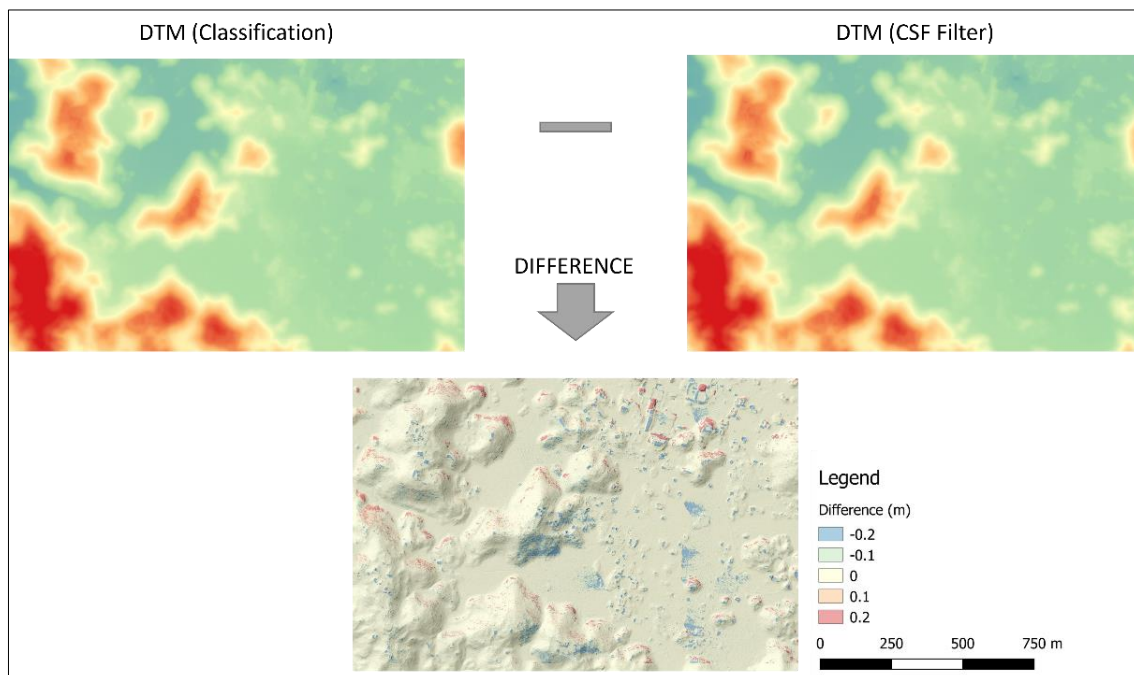

Figure 8: Differences between the DTM based on the ASPRS classification (a) and the CSF filter (b), shown on the hillshade model of the area of interest in Kiuic (c).

For the creation of the DTM there are several alternatives shown in the workflow of section 4.1. DTM's can only be derived from lidar data because of the multiple returns from which ground points can be identified, whereas Pléiades images do not provide any ground information in forested areas. To extract the ground points needed to produce a DTM from the lidar data, the first option is to classify the lidar point cloud by using the Cloth Simulation Filter (CSF). This filter method only needs a few easy-to-set integers and Boolean parameters, making it very straightforward to use. The method used by CSF is based on the cloth simulation developed by Weil in 1986. This involves turning the original point cloud upside down, dropping a virtual cloth on top of the surface where the final shape of the cloth is then used as a starting point for classifying the original point cloud into ground and non-ground

components [28]. The second option involves extracting the ground points from the classification that may already be present in the LAS file. By looking at the difference of the DTMs between the classification already given in the LAS files and the CSF Filter, small differences can be detected (Figure 8). On the one hand, positive values are located on the north side of the visible cone karst hills and on some relics in the northeast. On the other hand, negative values are localized on the south side of the hills and on quite a few Maya relics. The differences can be attributed to the various algorithms that were applied, each with their own strengths and limitations in specific terrains. Nevertheless, the use of morphological thresholds is a fundamental strategy for ground point filtering which produces different results [46].
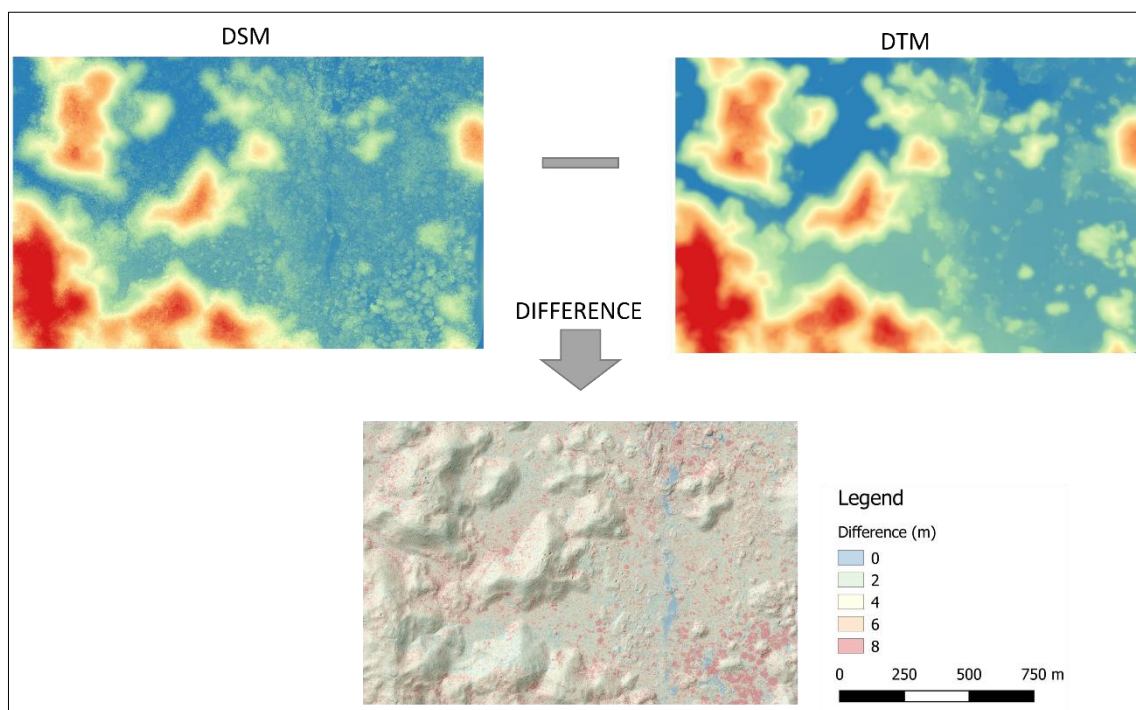

Figure 9: Differences between the DSM and the DTM from the G-LiHT lidar data, shown on the hillshade model of the area of interest in Kiuic.

The differences between the DSM and the DTM can also be visualized, which provides information about the height of the vegetation. Figure 9 shows that the height values between the two elevation models differ the most between the cone karst hills and the maya relicts where the highest vegetation occurs (Figure 9).

5.2   Visualizing an elevation model

The various DTM visualizations facilitate the detection of archaeological vestiges of the Mayan population. They allow conducting visual inspections to recognize geometric objects with shapes that are not found in nature (e.g., circles, rectangles, etc.). The presence of such structures, as seen on   Figure *10*, confirms that the Maya's may have once lived in this area of Mexico, the Puuc region. The visualization results can thus provide fundamental data to update the maps of the Maya relics or for purposes of project planning. Furthermore, it can also provide data for training deep learning models that can enable automatic detection in vast areas.
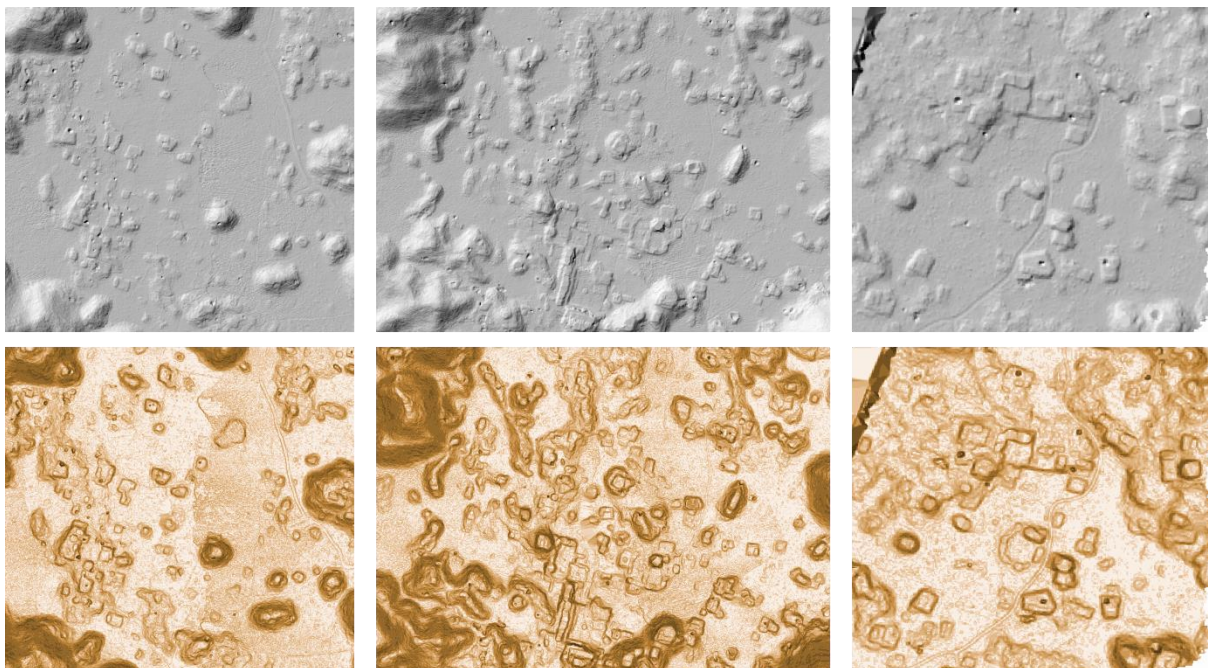

Figure 10: Visual inspection hillshade and slope visualization in ArcGIS Pro.

When analyzing data for your project, it is crucial to create multiple visualization models to ensure that all archaeological features can be identified. This is the case, for example, with hillshading because of its inability to represent linear objects that lie parallel to the direction of the light source and saturation of shadow areas [5]. A way to overcome this is to create more hillshades at different azimuths, which often reveals different features. This can be seen in Figure 11, where parts of the roads through the Maya relics may or may not be seen.
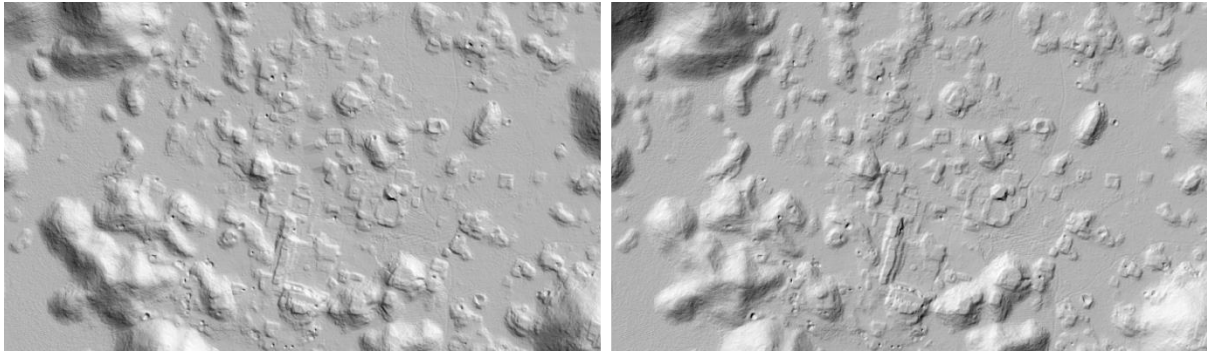
Figure 11: Two hillshades of the same landscape of Kiuic with the light source set from different angles (azimuth on 45 (a, left) and 315 degrees (b, right)) revealing parts of the way hidden to illumination from a single direction.

In the end, every visualization has its benefits and drawbacks to identify archaeological features. Some of them are better at picking out features in low relief topography, others at defining steeply sloping earthworks[18]. Many additional analytical models in function of preconceived project objectives can be manufactured with RVT (Figure 12).
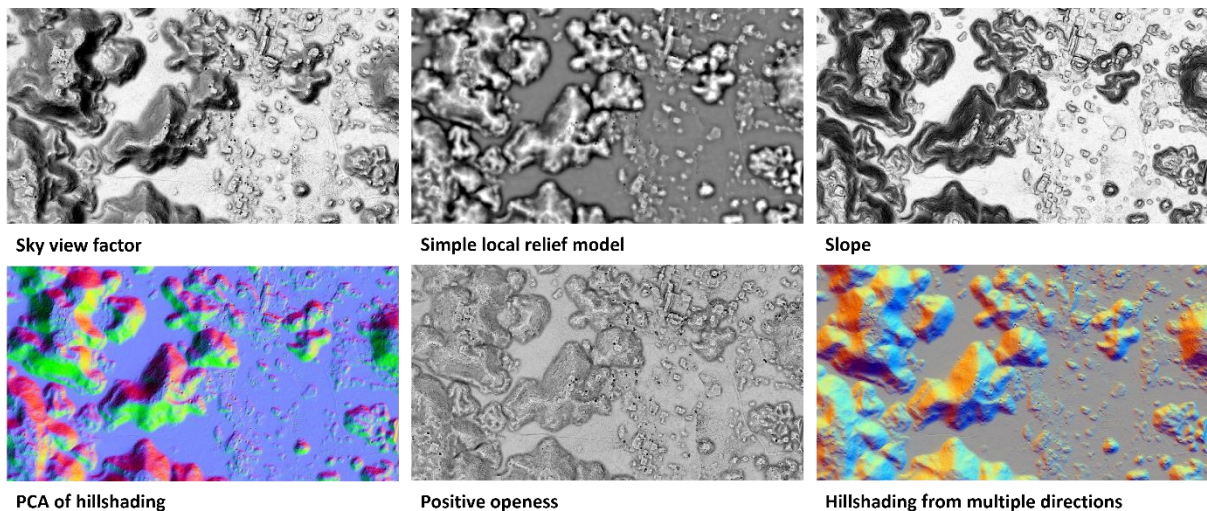

Sky view factor

Simple local relief model

Slope

PCA of hillshading

Positive openess

Hillshading from multiple directions

Figure 12: Kiuic analytical models made with the Relief Visualization Toolbox.

---

[18] Finding the most suitable visualization for a specific study area is often a trial and error story, more information about the advantages and disadvantages to identify archaeological features can be found in the relief visualization toolbox manual: https://www.zrc-sazu.si/sites/default/files/rvt_2.2.1_0.pdf

5.3   Orthophotos

Orthorectification is a geometric process that uses a relief model to eliminate the perspective effects on the ground by restoring the geometry of a vertical shot. The standard 3D model used for ground corrections of the orthophotos created by Airbus is the worldwide Reference 3D dataset[19], co-produced by Spot and the France's national survey and mapping agency. For the planimetric reset, the Elevation30 Ortho layer is used. This is a unique database derived from optical satellite data and merged with radar data. It is specifically designed for large area coverage, addressing mapping, GIS and military needs [47]. The correction of the altimetric reset uses an included DEM, depending on what is available: Reference3D DEM layer, SRTM or GLOBE [11]. For our case study, we produced two orthophoto's ourselves: one using the lidar height model and one using the Pléiades height model. A third orthophoto was obtained from Airbus. When comparing the three orthophoto's (Figure 13), we observed small shifts. Depending on the elevation model we used, small differences can be observed. Logically, a more accurate elevation model results in a better orthorectification. However, we can put the elevation model derived from lidar data that we used for orthorectification in this case study into question because there is a time difference of nine years between the lidar data and the Pléiades image.
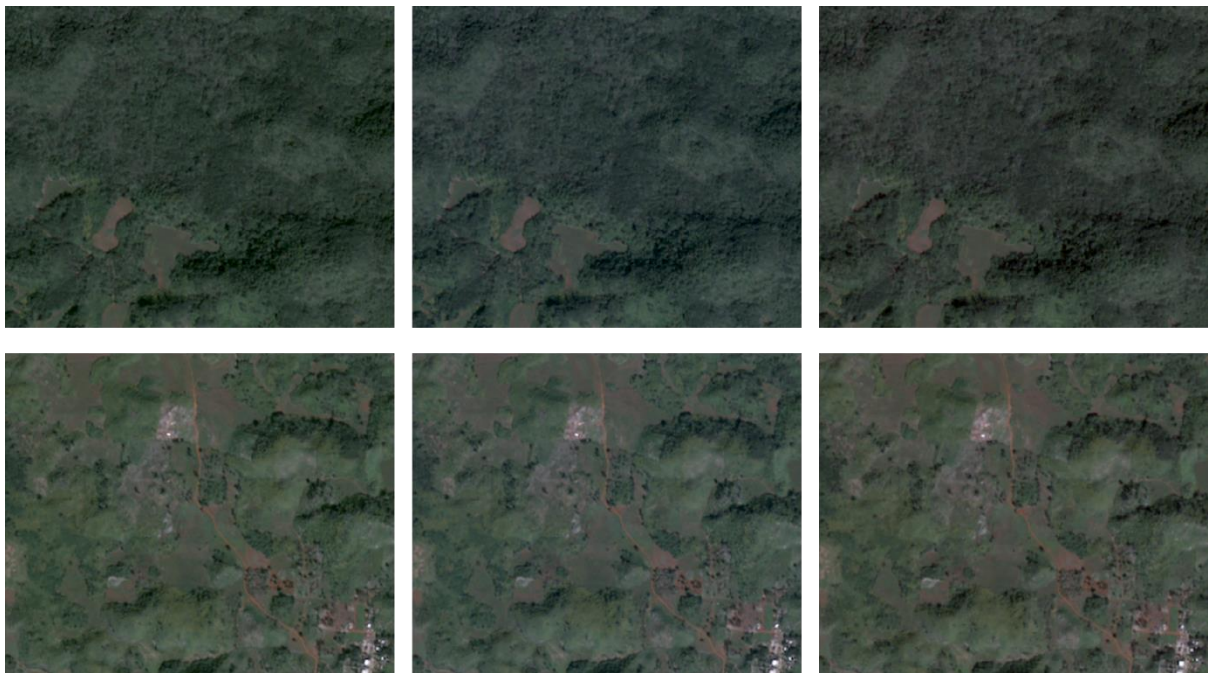

Figure 13: Ortho Images from Kiuic and Yaxhachén, made with DEMs of lidar (left), Pléiades images (middle) and the orthoimage delivered by Airbus (right).

---

[19] The document in the following link describes the specifications and format of the Reference 3D database: https://www.intelligence-airbusds.com/files/pmedia/public/r469_9_reference3d_product_description_201105.pdf

5.4 Pansharpening

As less details are visible on the multispectral image than on the panchromatic one, we applied a pan sharpening on the orthophotos described in the previous section. This results in a color image with a higher spatial resolution (Figure 14). Note that the terms 'high' and 'low' in the figure are used to describe the relative spatial resolution between images acquired by a single observation system.
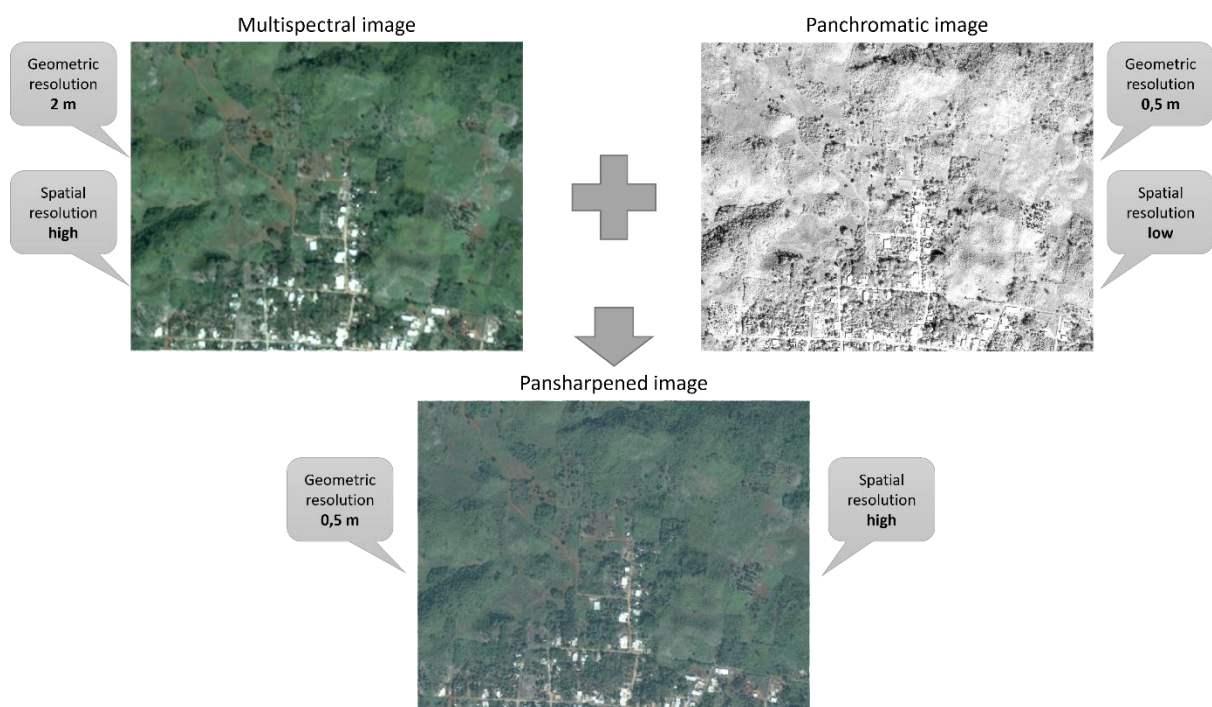

Figure 14: Method of the pansharpening using a multispectral and a panchromatic image.

5.5 False color composite

The four-band multispectral sensor (blue, green, red and near-infrared) from the Pléiades constellation does not provide as much opportunity for creating false color composites as would be the case for some other multispectral sensors (e.g. Worldview 2-3) or a hyperspectral sensor. However, it is possible to create the frequently used false-color image in which the reflectance measurements of the near-infrared band are used. Here one assigns the blue color to the reflectance's in the green light, the green color to the reflectance's in the red light and the red color to the reflectance's in the near infrared (Figure 15). This visualization gives vegetation a distinct red color, allowing it to be more easily distinguished from the surroundings.
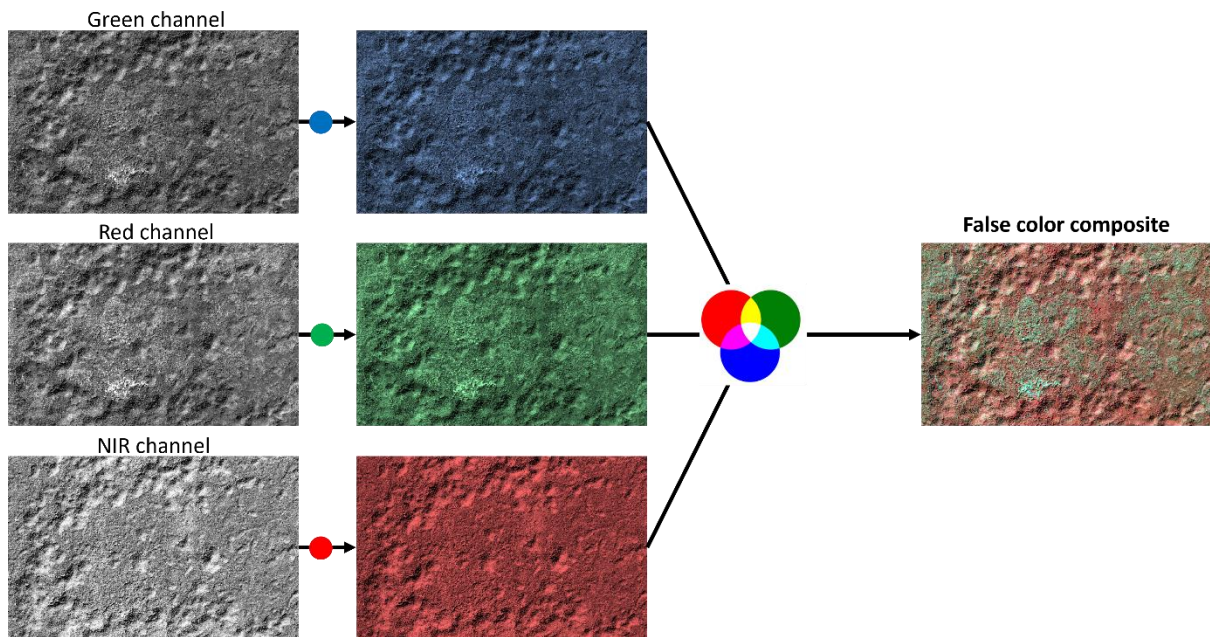
Figure 15: Method for the creation of the NIR False color composite with the Kiuic data.

By looking at different areas and comparing them to the true color image, we notice that the false-color image offers more details (Figure 16). It is easier to distinguish the vegetation from its surroundings because of the high reflectance of plants in the near-infrared spectrum. Exploiting this property, archaeologists today seek to identify vegetation patterns that are useful to locate archaeological sites in the forested Maya region. If vegetation is subject to some form of stress that interrupts its normal growth and productivity, it can decrease the chlorophyll production, resulting in less chlorophyll absorption in the blue and red bands [2].

Figure 16: False color composite images on the hillshade visualization below the true color image: the red indicates an abundance of vegetation (right), the image at the edge of the city with houses and streets displays much less vegetation (middle) and the center of Kiuic displays different vegetation patterns (left).
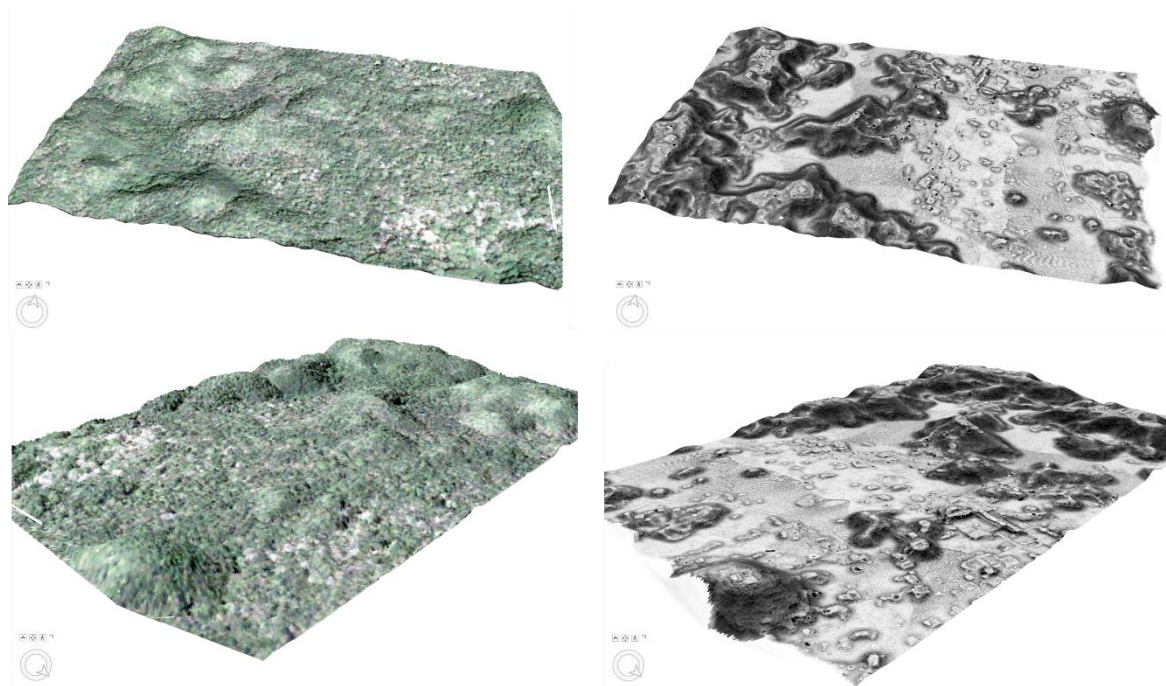
5.6   Basemap on a DEM

Figure 17: Different views of two different kind of visualizations for the place of Kiuic: the orthophoto from the Pléiades images (a, left) and the sky view factor (b, right).

Any raster layer or image can be draped on an elevation model. In the image below (Figure 17), we placed the orthophoto from the Pléiades images and a sky view factor visualization on respectively a DSM and a DTM of the area of Kiuic, where several Maya artifacts have already been discovered. The orthophoto unfortunately does not show much of the relics due to the vegetation cover. In contrast, the sky view factor visualized on the DTM of Kiuic provides a very good visual result, showing the relics in relief. The same process can be repeated with all the different visualizations described in section 5.2.
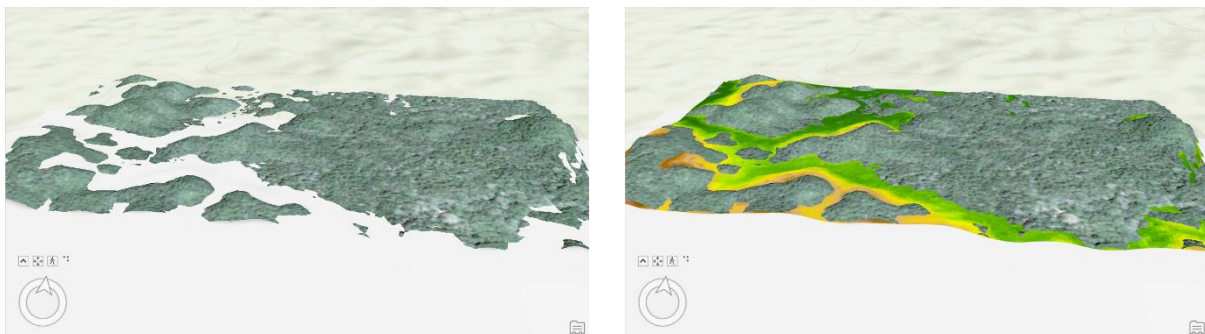

Figure 18: Importance of the cartographic offset to avoid gaps in the created 3D visualization: incomplete 3D model with gaps (a, left) and the same model filled with the DSM (b, right).

When your spatial data need to be displayed in 3D, you need to set height options and so it is important to select a vertical coordinate system. If the system of the data source is defined, the elevation unit of the layer is the same as the linear unit of the vertical coordinate system. To drape any visualization or orthophoto over a digital elevation model, one of the main things that can go wrong is the cartographic offset. This parameter vertically adjusts the z-value of the entire elevation layer and can be applied to adjust where features are drown vertically relative to the base height. Base heights determine the lowest elevation of features in the layer [44]. Depending on the value that was given to the cartographic offset, gaps or open spaces in the 3D model may be created in the visualization (Figure 18). To solve this problem, the parameter of the cartographic offset must be adjusted. By means of trial and error, you raise or lower all features in the selected layer by a given height [44]. In this way, the cartographic offset is optimized to make sure that the features are drawn just above the elevation surface of the DEM.
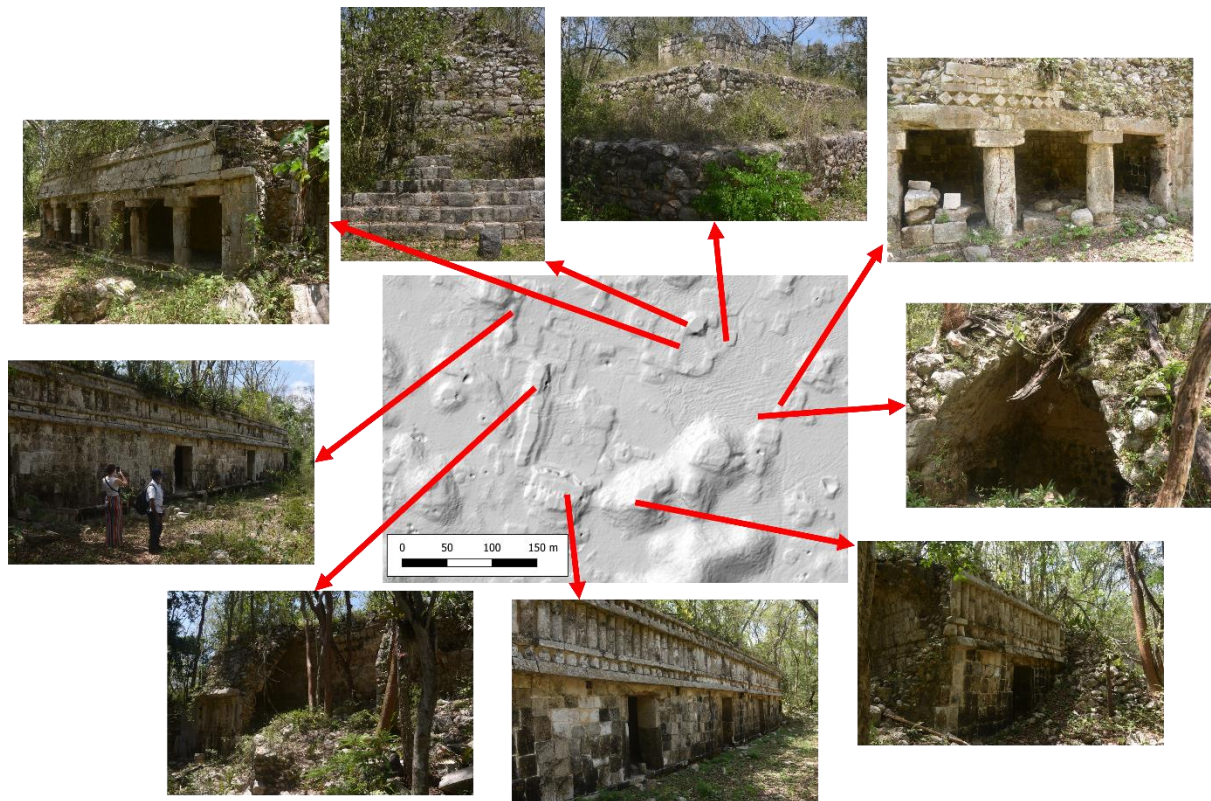
6 TERRAIN WORK

Figure 19: The various Maya remains seen during the fieldwork in Kiuic linked to the artefacts visible on a hillshade visualization.

Since the visualizations showed many unnatural shapes underneath the vegetation, which we suspected to be Maya remains, a site verification of Kiuic took place. On the first of April 2022 in the morning, with the help of INAH Yucatan, the area was explored and numerous large structures were observed. These remains, which were captured on photo, were linked to the visible structures on a hillshade visualization (Figure 19), confirming that what was observed on the terrain models did indeed match with the knowledge that Maya vestiges are present in this area. As a result, these types of visualizations can provide an enormous amount of information. Places where the archaeological potential is threatened by current activities (e.g. construction of the Maya train) can be mapped in this way.

7   RECOMMENDATIONS

There are some lessons to be learned from our case study:

- There is free lidar data online from which useful results can be obtained, such as the G-LiHT lidar data from NASA. Although everyone is always looking for the best resolution, these data turned out to be good enough to identify major archaeological features. More recent and qualitatively better lidar data exist but these are currently under embargo [22].

- The freely available lidar data were not recorded for the purpose of archaeology. However, this did not stand in the way of obtaining interesting results.

- The time gap between the lidar data and the Pléiades imagery did not matter much for this case study. Research related to vegetation would have been more difficult, but Maya vestiges do not change that much. However, it is very important to know the specifications of your data, and to be aware of these during the whole project.

- Notwithstanding the trend towards data sharing and open data, high quality geographic data with a high spatial resolution is often not readily available and free. Data can be quite expensive even, especially when they have to be newly acquired for a specific purpose or for a particular study area. This naturally puts limitations on research projects that can be conducted in specific areas, and it will also stand in the way of applying several techniques that were demonstrated in these guidelines. There is obviously no direct and general solution for this issue, but we hope that the case study presented in these guidelines can help convince funding agencies of the merits of using high resolution remote sensing data for archaeological research and prospection.

- Many different commercial software applications are available to process lidar data and stereoscopic imagery, with often more capabilities than the ones mentioned in this document. Rather than listing these applications here, we recommend an online search.

- Results are not always immediately what you expect. Take your time to understand and experiment with the different parameters. Good parameter settings are essential for good results and practice makes "perfect".

8 CONCLUSION

This assistance manual presents different possibilities for working with lidar data and Pléiades imagery in order to detect Maya settlements in Yucatan, Mexico. Airbus delivered ready-to-use Pléiades products, which were easily integrated in a GIS and which were combined with other geographic data such as the lidar point clouds of G-LiHT, released by NASA. Many visualizations have been made, with each workflow explained step by step. Depending on the objective, some visualizations yielded better results than others, especially regarding the recognition of archaeological relics in different kinds of landscapes.

The aim of the project drives us eventually to the question whether this all is interesting for archaeologists in Yucatan or not. The effectiveness in detecting Maya settlements in the Puuc region with remote sensing data became clear during a fieldtrip, where many archaeological Mayan vestiges could be seen. Photographs were taken of the various structures in Kiuic, which were then localized and linked with the hillshade visualization. Processing these kind of visualizations with lidar data has the advantage that large areas can be investigated in a short period of time in order to carry out more targeted field research. For example, this technology is already being used for the construction of the Maya train, one of the electoral promises of Andrés Manuel López Obrador, the President of Mexico since 2018. Since this is also being used in this region and research is being carried out with it (e.g. point cloud filtering), this manual may well be of added value to local archaeologists without skills in processing earth observation data.

9 REFERENCES

1. Panagiotakis, E., Chrysoulakis, N., Charalampopoulou, V., & Poursanidis, D. (2018). Validation of Pleiades Tri-Stereo DSM in urban areas. *ISPRS International Journal of Geo-Information*, *7*(3), 118.

2. Garrison, T. G., Houston, S. D., Golden, C., Inomata, T., Nelson, Z., & Munson, J. (2008). Evaluating the use of IKONOS satellite imagery in lowland Maya settlement archaeology. *Journal of Archaeological Science*, *35*(10), 2770-2777.

3. Hixson, D. R. (2013). The use of multispectral imagery and airborne synthetic aperture radar for the detection of archaeological sites and features in the western Maya wetlands of Chunchucmil, Yucatan, Mexico. In *Mapping Archaeological Landscapes from Space* (pp. 133-144). Springer, New York, NY.

4. Stichelbaut, B. (2018). Sporen van oorlog: archeologie van de Eerste Wereldoorlog. *Hannibal*.

5. Davis, O. (2012). Processing and working with LiDAR data in ArcGIS: a practical guide for archaeologists.

6. Werbrouck, I., Antrop, M., Van Eetvelde, V., Stal, C., De Maeyer, P., Bats, M., ... & Zwertvaegher, A. (2011). Digital Elevation Model generation for historical landscape analysis based on LiDAR data, a case study in Flanders (Belgium). *Expert Systems with Applications*, *38*(7), 8178-8185.

7. Janssen, L. (2017). *Infosessie: Brondata DHMV* [PowerPoint slides]. Informatie.vlaanderen@vlaanderen.be. https://docplayer.nl/51052796-Infosessie-brondata-dhmv.html

8. Drosos, V., & Farmakis, D. (2006). Airborne laser scanning and DTM generation. *Management and Development of Mountainous and Island Areas*, 206.

9. Environment Agency Survey Open Data. (2016). *LIDAR Returns* [Figure]. Flickr. https://www.flickr.com/photos/environmentagencyopensurveydata/29670476293

10. Inomata, T., Triadan, D., Vázquez López, V. A., Fernandez-Diaz, J. C., Omori, T., Méndez Bauer, M. B., ... & Nasu, H. (2020). Monumental architecture at Aguada Fénix and the rise of Maya civilization. *Nature*, 582(7813), 530-533.

11. Pléiades Imagery User Guide, ASTRIUM, Airbus Defence and Space. Available online: https://www.intelligence-airbusds.com/en/8718-user-guides#pleiades (accessed on 15 April 2022).

12. Panagiotakis, E., Chrysoulakis, N., Charalampopoulou, V., & Poursanidis, D. (2018). Validation of Pleiades Tri-Stereo DSM in urban areas. *ISPRS International Journal of Geo-Information*, *7*(3), 118.

13. Kpalma, K., El-Mezouar, M. C., & Taleb, N. (2014). Recent trends in satellite image pan-sharpening techniques. In *1st International Conference on Electrical, Electronic and Computing Engineering.*

14. Porta, E. W. (2005). John Lloyd Stephens, Frederick Catherwood and the Collaboration of Art and Science in the Rediscovery of Maya Civilization. *Terrae Incognitae, 37*(1), 63-74.

15. Rhyne, C. S. (2008). The Puuc Region. https://www.reed.edu/uxmal/Puuc.html

16. Cook, B. (2013). Metadata for NASA Goddard's lidar, hyperspectral land thermal (G-LiHT) airborne imager [pdf document]. bruce.cook@gsfc.nasa.gov

17. Schroder, W., Murtha, T., Golden, C., Hernández, A. A., Scherer, A., Morell-Hart, S., ... & Brown, M. (2020). The lowland Maya settlement landscape: Environmental LiDAR and ecology. *Journal of Archaeological Science: Reports, 33*, 102543.

18. Cook, B. D., Corp, L. A., Nelson, R. F., Middleton, E. M., Morton, D. C., McCorkel, J. T., ... & Montesano, P. M. (2013). NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager. *Remote Sensing, 5*(8), 4045-4066.

19. ASPRS. (2011). LAS specification. Microsoft Word - LAS 1 4 - R13.docx (asprs.org)

20. Esri. (2016). How To: Work with raster products in ArcGIS for Desktop, an introduction. https://support.esri.com/en/technical-article/000012290

21. Fernandez-Diaz, J.C., Carter, W.E., Glennie, C., Shrestha, R.L., Pan, Z., Ekhtari, N., Singhania, A., Hauser, D., Sartori, M. (2016). Capability Assessment and Performance Metrics for the Titan Multispectral Mapping Lidar. *Remote Sens*ing, 8(11), 936. https://doi.org/10.3390/rs8110936.

22. Ringle W.M., Gallareta Negron T., May Ciau R., Seligson K.E., Fernandez-Diaz J.C., Ortegon Zapata D. (2021). Lidar survey of ancient Maya settlement in the Puuc region of Yucatan, Mexico. PLoS ONE 16(4), e0249314. https://doi.org/10.1371/journal.pone.0249314

23. Palaseanu-Lovejoy, M., Bisson, M., Spinetti, C., Buongiorno, M. F., Alexandrov, O., & Cecere, T. (2019). High-resolution and accurate topography reconstruction of Mount Etna from pleiades satellite data. *Remote Sensing, 11*(24), 2983.

24. PCI Geomatics. (2022). Catalyst Professional. https://catalyst.earth/products/catalyst-pro/

25. CloudCompare. (2022). CloudCompare home. https://www.danielgm.net/cc/

26. Esri. (2022). ArcGIS Pro. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview

27. PCI Geomatics. (2020). DEM Extraction - Satellite – Geomatica 2018 & Older. https://support.pcigeomatics.com/hc/en-us/articles/203482839-DEM-Extraction-Satellite-Geomatica-2018-Older

28. Peng. (2016). CSF (Plugin). https://www.cloudcompare.org/doc/wiki/index.php?title=CSF_(plugin)

29. Esri. (2022). Create LAS dataset. https://pro.arcgis.com/en/pro-app/2.8/help/data/las-dataset/create-a-las-datasets.htm

30. Esri. (2022). LAS Dataset To Raster (Conversion). https://pro.arcgis.com/en/pro-app/2.8/tool-reference/conversion/las-dataset-to-raster.htm

31. Esri. (2022). Raster Calculator (Spatial Analyst). https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/raster-calculator.htm

32. Zakšek, K., Oštir, K., & Kokalj, Ž. (2011). Sky-view factor as a relief visualization technique. *Remote sensing*, *3*(2), 398-415.

33. Kokalj, Ž., & Hesse, R. (2017). Airborne laser scanning raster data visualization. A guide to good practice.

34. Challis, K., Forlin, P., & Kincey, M. (2011). A generic toolkit for the visualization of archaeological features on airborne LiDAR elevation data. *Archaeological Prospection*, *18*(4), 279-289.

35. ZRC SAZU. (2019). Manual: Relief Visualization Toolbox, ver. 2.2.1. https://www.zrc-sazu.si/sites/default/files/rvt_2.2.1_0.pdf

36. Esri. (2022). How HillShade works. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/3d-analyst/how-hillshade-works.htm

37. Esri. (2022). Create Ortho Corrected Raster Dataset (Data Management). https://pro.arcgis.com/en/pro-app/2.8/tool-reference/data-management/create-ortho-corrected-raster-dataset.htm

38. PCI Geomatics. (2020). Optical Satellite Orthorectification- Geomatica 2018 & Older. https://support.pcigeomatics.com/hc/en-us/articles/209885893-Satellite-Orthorectification-Workflow#h_529682713361535391309828

39. Esri. (2022). Pansharpening function. https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/pansharpening-function.htm

40. PCI Geomatics. (2020). Pleiades Pansharpening and Orthorectification. https://support.pcigeomatics.com/hc/en-us/articles/204186465-Pleiades-Pansharpening-and-Orthorectification

41. Humboldt State University. (2014). Natural and False Color Composites. http://gsp.humboldt.edu/olm/Courses/GSP_216/lessons/composites.html#:~:text=False%20color%20images%20are%20a,(i.e.%20near%2Dinfrared)

42. ESA. (2022). SNAP. https://step.esa.int/main/toolboxes/snap/

43. Esri. (2022). How To: Drape a basemap over LiDAR data in ArcGIS Pro Scene layer. https://support.esri.com/en/technical-article/000019199

44. Esri. (2022). Define height characteristics for layers. https://pro.arcgis.com/en/pro-app/latest/help/mapping/layer-properties/define-height-characteristics-for-layers.htm

45. NASA. (2022). The Stereo Pipeline. https://ti.arc.nasa.gov/tech/asr/groups/intelligent-robotics/ngt/stereo/

46. Chen, Z., Gao, B., & Devereux, B. (2017). State-of-the-art: DTM generation using airborne LIDAR data. *Sensors, 17*(1), 150.

47. Esri. (2022). Elevation 30. https://www.esri.com/partners/airbus-intelligence-a2T70000000TNKlEAO/elevation-30-a2d70000000mrbDAAQ