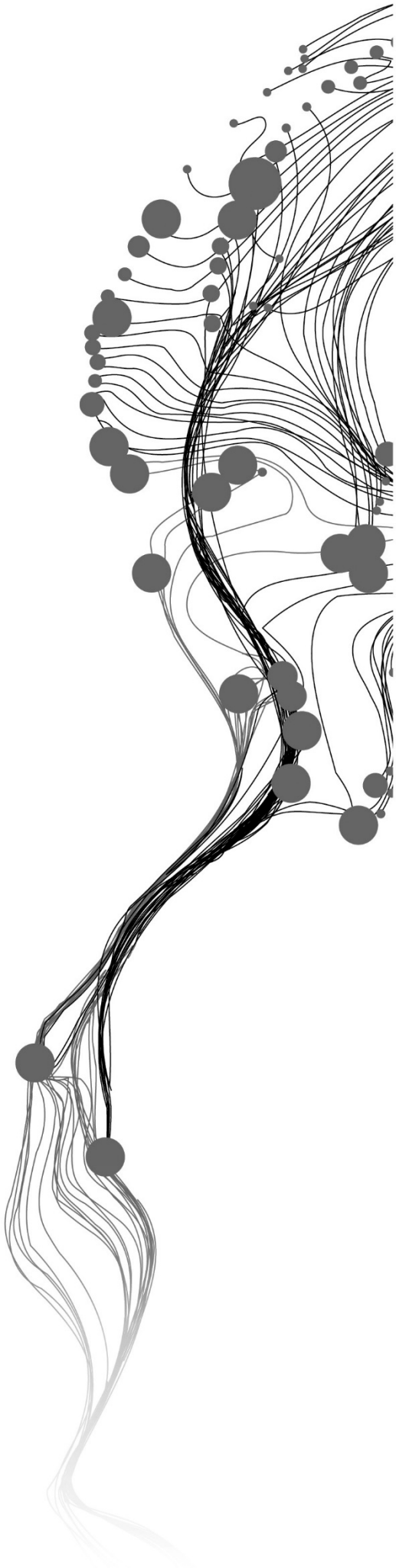
UAS BASED REMOTE SENSING TO DETECT FREE-RANGING ANIMALS IN A LOW CONTRAST, EXTREME TOPOGRAPHY HIGH DESERT LANDSCAPE

SILLA OTIENO AFWAMBA

February 2019

SUPERVISORS:

Dr. P. Nyktas
Dr. A.G. Toxopeus



UAS BASED REMOTE SENSING TO DETECT FREE-RANGING ANIMALS IN A LOW CONTRAST, EXTREME TOPOGRAPHY HIGH DESERT LANDSCAPE

SILLA OTIENO AFWAMBA

Enschede, The Netherlands, February 2019

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfillment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Natural Resources Management

SUPERVISORS:
Dr. P. Nyktas
Dr. A.G. Toxopeus

THESIS ASSESSMENT BOARD:
Dr, Yousif A., Hussin (Chair)
Dr, Francesco Pietro Fava (External Examiner, International Livestock Research Institute (ILRI, Kenya)

ABSTRACT

Unmanned Aerial System presents a new frontier in emerging technologies for the management of rangelands. Although satellite imagery has been used extensively in abundance estimation and enumeration of animals such as birds and large herbivores, in rangelands, use of UAS data is still on a small scale covering small areas. UAS data, however, offers several advantages over traditional ground-based methods because of its flexibility in animal detection.

This study was done in Lefka Ori Mountains in the Island of Crete, Greece. It aimed to investigate the possibilities of using UAS data in detecting and counting animals (goats and sheep) in a low contrast, extreme topography high desert environment. It examined the possibilities of using motion image, point cloud, and orthophoto image by applying manual and semi-automatic techniques using only RGB images on one side and fusion of RGB images with height information approach. Also, the motion images were georeferenced and used for generating an estimate for animal presence and location. Short time interval image differencing was done on still image pairs and analyzed to generate true positives, false positives and false negatives.

UAS data collection entailed acquiring high-resolution images concurrently with ground counts of animals. The accuracy of the methods described was done using RMSE, Precision, Recall, and F measure. Motion image georeferencing had an RMSE of 15m. The manual counts had an overall accuracy of 89% for image differenced results, but for original images, the accuracy of the count was 11%. For a semi-automated method, the overall score using fusion approach for the selected test sites was 56% recall with a 75% precision and an F-score of 64%. The overall score for UAS RGB Image approach for small test sites was 50% recall rate with a precision of 7% and an F score of 12%. For the large site (0.9ha), fusion approach had 47% recall rate achieved with a 1% precision and an F-score of 2 % while RGB image approach did not have any results.

The results showed that small test sites within the large test site demonstrated better results compared to the large site which demonstrated poor results. It also showed feasibility for small surveys however, a combination of manual and semi-automatic counting can be feasible for large sites. Manual counts from both on the small and large area were high therefore also feasible. Fusion approach had better results than using UAS RGB images only approach.

Keywords: Image Differencing, Unmanned Aerial System, Image Fusion, Precision, Recall, Rangeland, High Desert, Normalised DSM, Thresholding, RGB images.

ACKNOWLEDGMENTS

I give thanks to God for enabling me to accomplish this MSc. I would also like to thank the NUFFIC and the ITC faculty of the University of Twente for giving me the opportunity to complete this Msc through an NFP/OKP scholarship. My special gratitude also goes to Dr. Panagiotis Nyktas and Dr. A.G. Toxopeus Bert, my first and second supervisors respectively, for their valuable advice, support and patience. My gratitude also goes to Timothy Roberts, the drone expert, for his valuable assistance and innovation during fieldwork planning and execution.

Special gratitude also goes to Samaria National Park management body employees and the president, Petros Lymberakis, for their support in providing valuable information on Lefka Ori Mountain, permissions as well as the vehicle we used during field work.

I would also like to thank NRM staffs for their teaching and support. Thanks to all my NRM classmates of 2017/2019 for their support during my study. Special gratitude goes to my class mates; Isaac Oliech and Max Korir for their support during the fun and difficult field work moments in Crete. Lastly, I would like to thank my family for their love, prayers and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF APPENDICES

LIST OF ACRONYMS

AAT	Automatic Aerial Triangulation
BBA	Bundle Block Adjustment
DSM	Digital Surface Model
DTM	Digital Terrain Model
ENM	Environmental Niche Models
FMV	Full Motion Video
FN	False Negatives
FOV	Field of View
FP	False Positives
GNSS	Global Navigation Satellite System
GPU	Graphics Processing Unit
ICA	Independent Component Analysis
INS	Inertial Navigation Systems
LiDAR	Light Detection and Ranging
MISB	Motion Imagery Standards Board
nDSM	Normalised Digital Surface Model
PCA	Principal Component Analysis
TP	True Positives
UAS	Unmanned Aerial Systems
UAV	Unmanned Aerial Vehicles
UgCS	Universal Ground Control Station
UTC	Universal Coordinated Time
VHR	Very High Resolution
D-RTK-GNSS	Differential Real Time Kinematic Global Navigation Satellite System

1. INTRODUCTION

1.1. Introduction

Rangelands cover about 50% of the land cover in the world, and the geographic extent including the resources of rangelands make their sustainable use and management very significant (Menke & Eric Bradford 1992). Rangelands are characterized by grasslands, shrublands, woodlands and deserts and these are known by different names such as savannas, chaparrals, steppes, prairies, tundras and woodlands depending on the global location. Rangeland management involves planning and administration for maximum benefit especially in livestock production but with sustainable use of the resources in focus. The core components vital for the management of rangeland include an inventory of both plant and animal species, monitoring data on soils as well as vegetation. According to Menke & Eric Bradford, (1992), proper use of rangelands for economic gain through livestock production; management of wildlife for recreational use; and conservation of biological diversity is the key to efficient management of rangelands.

These arid and semi-arid environments are characterized by low average rainfall which influences the abundance of flora and fauna in these areas as well as human activities. High deserts are mostly located on high altitude ranges and are characterized by hot and cold temperatures, dry summers and low rainfall (Laity 2009). Examples include Gobi Desert in Mongolia, Lefka Ori mountains in Crete and Oregon high desert in the USA. Some of the high deserts such as Lefka Ori mountains (White Mountains) in Crete are home to endemic flora and its important floristic diversity is well known internationally and documented. (Fernández-Calzado et al. 2013). This high plant diversity offers an opportunity for the study on species richness and spread.

Efficient management requires correct parameters such as animal population; species; topography; weather and water sources distribution but of importance is animal population. Grazing is a major rangeland land use activity that determines the health of rangelands and is of key concern in rangeland management. According to Shroder, Sivanpillai, Angerer, Fox, & Wolfe, (2016) overgrazing occurs when vegetation material is repeatedly removed by livestock and other wild herbivores which exceeds the vegetation ability to regrow. He further stated that continuous overgrazing over a long period could lead to rangeland degradation. Having a clear understanding of the long term effects of grazing on rangelands is essential in the investigation and assessment of the ecosystem services for a rangeland system(Hao et al. 2018). However, data for generation of parameters useful for assessing spatial grazing patterns and distribution are either rare or inaccurate. Furthermore, existing tools for acquisition of parameters for rangeland management faces many challenges which limit accuracy due to the geographic extent of the ranges (Gonzalez et al. 2016).

Animal abundance information is essential for successful management of the rangelands. Assessing spatial distribution and grazing patterns requires accurate data on animal population and this is done either through direct or indirect counting using ground survey or remote sensing(Van Gemert et al. 2015). In-situ or ground survey and remote sensing methods have been used to assess the rangelands and provide a guide to good management practices. According to Rango et al. (2009) Government authorities and private authorities regularly provide inventories and assessment of rangelands under their jurisdiction, and due to the vast nature of rangelands, often ground-based survey alone is inadequate. The constraint is the large sample sizes required; time required; cost and accessibility in remote areas. Sometimes official

statistics generated are inaccurate and without reliable numbers, it is difficult to create solutions for managing the rangelands (Vogiatzakis & Rackham 2008).

Remotely sensed imagery, however, has several benefits over ground-based surveys due to their large spatial coverage, and accessibility to difficult areas as well as faster processing of data(Gonzalez et al. 2016). It also acquires data at a point in time and can be stored and referred to in future. The most commonly used methods for acquisition of imagery are satellite imagery, aerial photos, and Unmanned Aircraft Systems (UAS).

The spatial and spectral resolution of these images is dictated by the sensor and platform used, and it ranges from low to high resolution. High spatial resolution imagery of 0.4m, 0.5m, 3m as well as 5m are currently available from various satellite image providers. Yang et al.(2014) used satellite imagery of 0.5m resolution to automatically detect Wildebeest, Zebra and Gazelle in Maasai Mara National Reserve, Kenya. The result had an omission error of 8.2% and commission error of 6.6%. Also, (LaRue et al. 2015)In their study on testing methods for using high-resolution satellite imagery to monitor polar bear abundance and distribution used a 0.5m satellite imagery to count Polar bears in Rowley Island in Canada and had an 87 % accuracy. Satellite data, however, is limited by the spatial resolution and features smaller than the spatial resolution cannot be easily detected. Large animals can be detected easily compared to small animals.

Majority of researches done in animal detection have employed aerial imagery in rangeland applications using 25cm resolution (Rango et al. 2009). Aerial imagery using aircraft has some advantages over satellite imagery in that its images can be tasked at a higher spatial resolution (Hollings et al. 2018). Terletzky & Ramsey, (2014) detected and counted horses and cattle within rangeland in Utah, the USA using aerial images of 25cm and the result had an 82% accuracy. (McNeill et al. 2011) was able to get a detection rate of between 90-95% using aerial imagery of 0.5m resolution in their study on Adelle Penguins in Antarctica. However, the limitations of aerial imagery include landing zones during flight operations and weather conditions. It may thus not be very suitable for detecting free ranging medium-sized animals in an extreme environment such as a high desert characterized by a rugged landscape and low contrast.

A resolution lower than 25cm is therefore needed to improve detectability. Increase in the use of UAS in the management of resources has been on an upward trend perhaps because of its very high spatial resolution (VHR); flexibility in that it can be deployed anywhere. Hollings et al., (2018) reported that its ability to capture images with a submeter resolution has been its greatest strength. It can also withstand small environmental effects, but large impacts such as strong winds are a challenge. However, technological developments in this sector promise to improve on these challenges (Rango & Laliberte 2010).

UAS were initially being used for military applications but are now becoming popular in many civilian applications. For this study, the term Unmanned Aircraft System (UAS) has been used to describe the entire components which include the UAV, the control station, that is the person controlling the aircraft and the software applications involved, and lastly the system that connects both (wireless link). Both high resolutions, still and motion images, are obtainable using UAS. The UAS is increasingly becoming standard platforms for photogrammetric data capture applications and are providing an alternative for traditional photogrammetric data capture, at a higher spatial and temporal resolution (Previtali et al. 2012). The UAV platforms include balloons, kites with the popular ones being rotary wing and fixed wing aircrafts. Both rotary and fixed wing have improved performance regarding endurance and wind resistance (Benassi et al. 2017). The system can be built up at reasonable prices, and therefore their use is cost-

effective. The types of cameras used for earth observation include RGB, Multi Hyperspectral, and LIDAR.

Even with some limitations such as low sensor performance, low payload endurance, instability of platform due to winds, regulations and insurance, the use of UAV has become popular for earth observation due to its flexibility and cost benefits as compared to the manned aircraft. Moreover, terrestrial and aerial techniques can be applied in data processing 2D and 3D information. This has been motivated by the rapid development of automated algorithms and procedures for image processing in computer vision such as Structure from Motion (SfM), Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Random sample consensus (RANSAC) among others.

There are some notable examples of the use of UAS in rangeland and wildlife management. Sykora-Bodie et al. (2017) used UAV in their study on assessments of sea turtle populations in coastal marine ecosystems and were able to identify the sea turtles from the UAV images through visual interpretation. Hodgson et al., (2018) did a study on identifying and counting seabird colonies and concluded that UAV images had a 43-96% more accuracy than a ground survey. A study by Seymour et al. (2017) on marine seals population using thermal imagery acquired from UAS generated a 95-98% accuracy when auto classified compared to manual counting. In rangeland management, Rango & Laliberte, (2010) used UAV images to asses rangelands, especially remote sites. Generally, most research done has mostly focused on high contrast small areas where features are easily discernible but little research has been done on rangeland management using UAV focusing on low contrast areas with medium sized free grazing animals.

Many methods have been explored and tested on their applicability in directly counting animals. Generally, the focus has been on large animals which are easily detectable on RS images. Manual methods have been the most used and tested, however recently advancement in machine and deep learning methods has enabled automatic and semi-automatic methods to be used. These have been tried and tested on their applicability albeit on a small scale (Hollings et al. 2018). The methods have also generated different results often dictated by the type of imagery, size of the study area and the animal numbers within the area. According to (Terletzky & Ramsey 2014) animals which are closer together such as sheep tend to reduce the accuracy of the counts. Techniques such as supervised classification, image segmentation, object-based image analysis, unsupervised classification, use of algorithms and neural networks have been applied.

Mostly focus has been on high-resolution satellite imagery, but some methods have been employed on VHR from UAV. Yang et al.,(2014) used artificial neural networks to identify wildebeest, gazelle, and zebras in Masai Mara National Reserve from VHR imagery. Groom et al. (2013) also used aerial imagery of 8cm spatial resolution to detect Lesser flamingoes using OBIA and had a success rate of 99% detection in comparison to manual interpretation. Change detection method such as image differencing was also used in a study to identify polar bears, and the success rate was 87% (LaRue et al. 2015) The following section describes the various methods that have been used in direct animal detection and counting.

1.1.1. Manual Counting

Direct counting of animals on images is has been the most commonly used method and the most studied before the introduction of the automated and semi-automated because it can easily correlate with the traditional ground methods (Terletzky & Ramsey 2014). Although its ideal for small areas mostly the method is labor and time intensive as well as relatively costly. Terletzky & Ramsey, (2014) further stated that the subjective nature of the method and homogeneity of some features for instance, crowded sheep and body size and background leads to overestimation and underestimation. Due to some element of subjectivity, manual method of detection and counting of animals generally are prone to errors especially

heterogenous and low contrast areas. Therefore, its limited to small regions when used alone. Counting animals from large volumes of images generated by drones is expensive and time-consuming.

1.1.2. Automatic and Semi-automatic techniques

Various deep learning techniques such as unsupervised and supervised learning have been developed for automated and semi-automated animal abundance estimation especially for large and inaccessible areas with low animal densities (LaRue et al. 2015). Also, deep learning architectures such as Artificial Neural Networks and Convolution Neural networks have also been tested (Van Gemert et al., 2015) although most of the studies undertaken were focused on relatively small areas within a homogenous environment and were proof of concept. Recent advancement in image processing software capabilities has improved both objects based and pixel-based approach in features identification (Rango & Laliberte 2010) the following section discusses various machine learning methods that a have been used in animal detection and counting.

Image differencing

Image differencing is an algebra change detection method commonly used in land use and land cover changes. This is a pixel-based method and requires the before and after image, datasets to be of the same pixel size and co-registered. Recent studies have also shown that this method has been successful in detecting animal movements using changes in spectral reflectance between two sets of images of the same place taken within a short period. The period could be 1 hour, or half a day or one day. These changes in animal movement are then counted to estimate animal population. (LaRue et al., 2015a; Terletzky & Ramsey, 2016).

Terletzky & Ramsey, (2014) used short time interval image differencing using principal components analysis on a 25cm resolution aerial image to estimate the population of horse and cattle in open pasture with high contrast. In their research on the polar bear population on Rowley Island in Canada, (LaRue et al. 2015)used image differencing method on a 50cm resolution satellite imagery and were able to detect 87% of the polar bears with less false positives. The mean commission error, however, was 53%. This could have been due to mismatches during image registration and alignment of two images of the same place. Also, the misidentification of some background objects was another limiting factor (Hollings et al. 2018)

Spectral Segmentation and thresholding

Image segmentation provides building blocks for object-based analysis. Spectral segmentation method uses spectral and texture information to classify images. One significant difference is that the set of points for differentiation are fixed before running as opposed to differentiation determined by the data. This is further enhanced through the use of filters to identify the spatial patterns (Yuan et al. 2014). Filters such as area size, compactness, maximum and minimum axis are used to segment the image. The critical issue is to minimize overlap. By partitioning an image into non-overlapping areas, objects are identified. According to (Yuan et al. 2014) high-resolution images have a rich texture which improves segmentation. Trathan, (2004) used image segmentation and thresholding methods to estimate the population of macaroni penguins using aerial images of 5cm resolution. The result had an R^2 of 0.95-0.98 and the manual counts from the image were 3-6% higher than automated counts.

Supervised and unsupervised classification

In this pixel-based method, image processing algorithms are used on trained data of known objects. The spectral signatures generated is then used to classify the whole image. (Hollings et al. 2018). Animals with similar spectral signatures based on the training data are classified and counted. LaRue et al. (2015) used this method to count polar bears in Rowley Island, Canada. However, the major limitation was a high number of false positives since some non-polar bears were also counted as polar bears. Supervised classification method, however, has not been as successful as other algorithm methods partly because it heavily depends on the user knowledge and experience in the identification of training data (LaRue et al. 2014). Unsupervised classification, on the other hand, performed better than supervised classification. This classification algorithm groups pixels together based on their spectral similarity. However, it had a high level of overestimation.

Object-based Image Analysis

According to (Yang et al. 2014) OBIA has better accuracy in comparison to pixel-based methods because it minimizes spectral overlap and can incorporate expert knowledge. However a study comparing OBIA and pixel-based in the detection of mammals in Maasai Mara, Kenya did not find any significant difference in performance of the methods for detection and counting(Yang et al. 2014). OBIA uses the object as the prime focus and the neighborhood of the object provides the context for defining the object based on texture and shape (Hodgson et al. 2018). A major limitation is the time it takes to process an image. According to Duro, Franklin, & Dubé, (2012), running certain algorithms can take up to 20 min per square kilometer.

The methods have had a higher success rate in identifying large animals and within a homogenous environment, but few types of research have been done on detection of, identifying and counting of medium-sized animals within open rangelands using very high-resolution UAV images. According to Hollings et al., (2018), the major drawback is the volume of images generated. Orthorectification is also a significant issue, especially when using methods such as image differencing and classification. However, using programmable Graphics Processing Unit (GPU) technology can improve accuracy (Jin et al. 2014).

1.2. Research problem

The animal detection and counting are influenced by the method used and the type of environment as well as the size of the area. Laliberte & Ripple, (2003) noted that differentiating animals from their environment in imagery largely depends on their contrast with the background (Figure 1.1) A low contrast environment presents a great challenge when using either manual or the semi-automated method in animal estimation and identification. Various change detection methods such as image differencing have been used especially to assess land cover and land use changes. Combining Image differencing and thresholding offers a better way of detecting animals using both manual counting semi-automatic counting. Recent studies have shown that this change detection method, short interval image differencing, can be used to estimate directly population of animals (Terletzky & Ramsey 2014).

Despite the increasing importance of UAV technology in animal abundance estimation, extensive assessment of the detectability rate and animal species detection errors in low contrast and irregular terrain environment has not been adequately assessed.

Figure 1.1: Low contrast environment part of the white mountain(left) showing animals in the background which cannot be easily differentiated with the background. The high contrast environment image(right) was used to explain the difference in the visual interpretation of the images.

Crete island landscape ranges from coastal lowlands to the high altitudes in the Lefka Ori mountains (white mountains), Psiloritis, and Dikty mountain ranges and is characterized by arid and semi-arid vegetation. Human settlements and tourism in West Crete have affected the Cretan natural landscape negatively due to potential economic, environmental and sociocultural impacts (Briassoulis 2003). Pastoral systems and free grazers is one of the major lands uses in West Crete and, free-grazing livestock especially goats and sheep on the low and high areas coupled with overgrazing has altered the original landscape and increased the area under semi-natural vegetation and subsequent degradation of the land. Due to the low contrast between the background environment and the animals coupled with extreme topography, ground observation is not a very reliable option over manual counts from raw images. (Figure 1.1). The animals not only have the same color and shape as the stones and vegetation but also show almost similar pixel intensity values. (Figure 1.2).

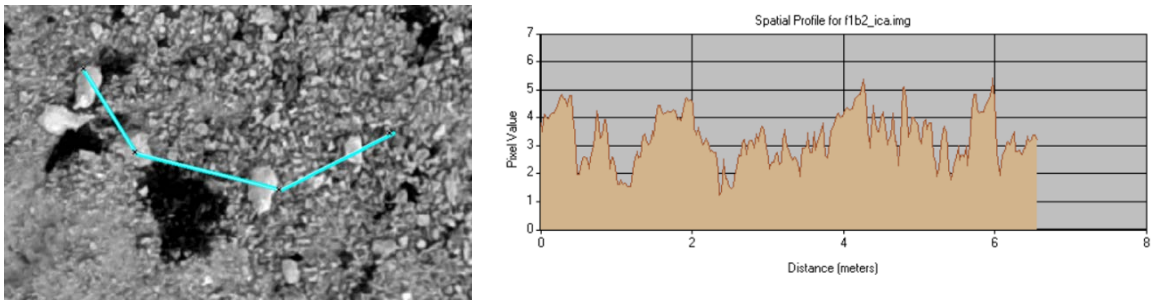


Figure 1.2. Spatial profile of a site having animals and background features. The 4 major peaks in the profile represent the features along the polyline drawn on the image. Feature 1 and 3 are animals while 2 and 4 are limestone rocks, but their spatial profile is almost similar.

The aim of this study, therefore, is to evaluate the feasibility of using UAS datasets in detection and counting of animals, specifically goats and sheep in the study area. The primary focus will be on direct detection on images acquired from a UAV platform using a change detection method. The study will attempt to use short interval image differencing using Independent Component Analysis coupled with thresholding in low contrast and heterogeneous area to assess the feasibility of the two methods in combination. It will also assess the feasibility of image fusion of UAS RGB images and height information from normalized DSM dataset to improve detection rate in image differencing.

1.3. Aim

To determine the applicability of UAS data in detecting and counting medium-sized free ranging animals for rangeland management.

1.3.1. Objectives

a. Assess the potential of using UAS as an alternative tool for the collection of animal presence data.
b. To test the feasibility of applying short time interval image differencing change detection method using independent component analysis on UAV images to detect animals.

1.3.2. Research questions

a. Can motion images provide high accuracy in spatial location and distribution of animals?
b. Can image differencing using Independent Component Analysis provide overall detection rates comparable to ground-based and manual survey in the detection of animals like goats and sheep in open rangeland?
c. Can fusion of RGB image and height information improve the detection of animals in open rangeland?
d. Are the accuracy rates high enough to consider the feasibility of the image differencing method?

2. MATERIALS AND METHODS

2.1. Study Area

Lefka Ori mountain is an extensive mountain range located in West Crete, Greece. The study area is located on geographic coordinates; 35.309413° and longitude 24.047754° as shown in the study area map on Figure 2.1. This area experiences a Mediterranean type of climate and annual precipitation reaches to more than 2000mm in the highest parts of the white mountains (Grove & Rackham 1993). The study area which is located on the white mountains is characterized by mainly limestone type of rocks thus giving the area a white to grey appearance. Generally, the area has a Mediterranean type of vegetation consisting of forest, maquie, phrygana and steppe (Vogiatzakis & Rackham 2008).

The main land use and land cover is rangelands and is primarily used for grazing of sheep and goats. Villages and hamlets mostly characterize settlements and land ownership is private in many areas in West Crete among landowning families although certain areas though have common land for all families. Transhumance used to be a major practice in this area by some farmers but it's now in its decline. (Lorent et al. 2009). During spring and summer between March and August, the animals are driven to the high mountains, but the sheep are attended to by herders while the goats are left to move on their own. Water points for the animals are located within the mountain range (Figure 2.2) and as winter approaches from September, the animals move towards the low lands of the mountain.

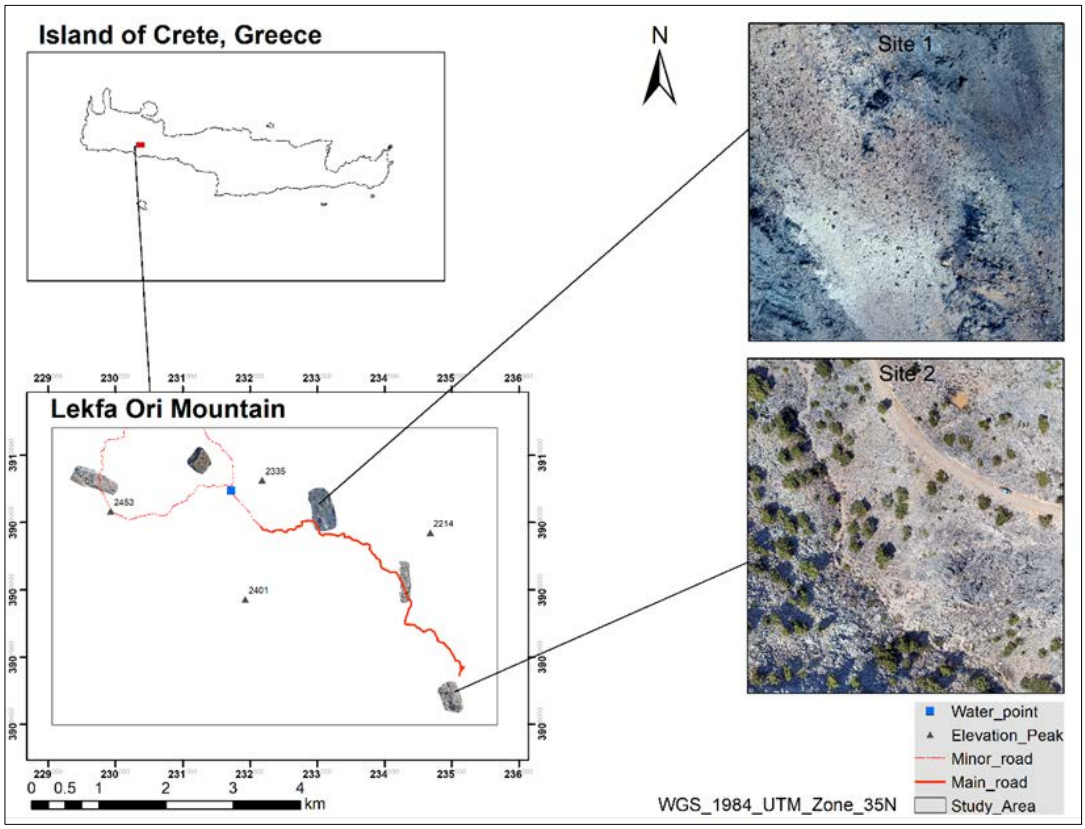


Figure 2.1: Map showing the study area. Location 1 and 2 were chosen for subsequent study due to the presence of animals in these areas.

Figure 2.2: Water points.

2.1.1. Field sampling design

Belt transect method was used to identify sampling frames for data collection as well as image acquisition. This was done using high-resolution imagery of the area acquired from google earth and topographical map of the area. For each transect, a site with animal presence was selected purposively to act as the control segment. The five transects of 0.6km by 0.2km were chosen based on their high suitability and probability of finding domesticated animals (Gormley et al. 2011). The areas were chosen because they were at and above the tree line with no trees to obstruct the detection. Secondly, animals in these areas were left completely to graze unmanaged. Figure 2.1 shows the sampling areas selected for the study. However, animal movement in this area is determined by the seasons and the survey was taken during the month of September when animals are moving towards the lowlands down from the white mountains due to the onset of the cold season. Out of the six sampling areas, only 2 transects were selected for further analysis because of the presence of the animals on these 2 sites. A further 4 small sites were created in site 1 with each having animals.

2.2. Data Collection

Phantom IV was used in imagery collection (Figure 2.3.). The phantom quadcopter is the most used UAV equipment in ecology and in other sectors because it requires little training and can be used by both professionals and semi-professionals (McEvoy et al. 2016). It can take off from and land vertically on a limited area; has onboard flight comptroller with a compass and inertial, gyroscopic, barometric and GNSS sensors. One of the key features is its 3-axis gimbal which enables the production of stable imagery. The summary of the specifications for the Phantom 4 UAV is provided in Table 2.1.

Table 2.1. Summary of Phantom 4 specifications.

Aircraft	**Description**
Weight	1388g
Obstacle Sensing Range	0-30m
Performance	
Maximum Flight time	30min
Operating Temperature	0-40°C
Navigation	GPS/GLONASS
Gimbal controllable range	Pitch: -90°-+30°
Gimbal Stabilization	3-axis (pitch, roll, yaw)
Max Ceiling above Sea Level	6000m
Camera Sensor	1" CMOS, Effective pixels: 20M
Camera Lens	FOV 84° 8.8 mm/24 mm
Vision System	Forward Vision System
	Backward Vision System
	Downward Vision System
Video Recording	MP4/MOV (AVC/H.264; HEVC/H.265)
Photo	JPEG, DNG (RAW), JPEG + DNG

2.2.1. UAV flight mission

Data acquired namely UAS RGB images, point cloud, DTM and motion images were used for this study. UAV imagery was collected between 11th September 2018 to 21st September 2018 from 10 am to 3 pm depending on the weather conditions such as cloud, sun intensity and wind speed. Test flights of 35, 60, 80, and 120m heights were done to determine the best height and speed regarding image quality and to avoid disturbance of the animal feeding process. Subsequently, 40m flight height with a GSD of 3cm was used. For each photogrammetric flight plan, an overlap of 80% and side-lap of 60% was used with a flight speed of 3m/s. Each flight mission covered approx. 5-9ha area.

Figure 2.3. Phantom 4 Unmanned Aerial Vehicle on the ground.

2.2.2. Universal Ground Control Station software (UgCS)

Lefka Ori mountain site has a rugged, steep and rocky terrain with limited visibility. Therefore, mission planning required a software that can accommodate these characteristics. In creating the flight plans, Universal Ground Control Station software was used. UgCS is a drone flight planning software which is compatible with many unmanned aircrafts. The software has built-in automated photogrammetric mission planning tools, and one of its key features is terrain following mode which enables the UAV to fly at a relatively constant altitude above the ground level thus limit effects of obstacles that can damage the drone as well as keep it in constant Ground Sampling Resolution (GSD) as shown in Figure 2.5.

Other key features include offline usage by utilizing the offline maps; flights can also be resumed from a certain point after a change of battery; Missions can also be planned and actualized even in a remote area such as a desert area without an internet connection. It not only has default DEM (SRTM) which can be used, but any DEM can be imported to the system to improve the accuracy of the mission process. Lastly, it has a telemetry window that allows the user to monitor the movement of the drone and it displays information such as battery charge; GNSS signal quality and link; altitude and current course, heading and speed (Figure 2.4.) *wind.com website* was also used to predict wind patterns and identify days when the wind is favorable for flight missions.

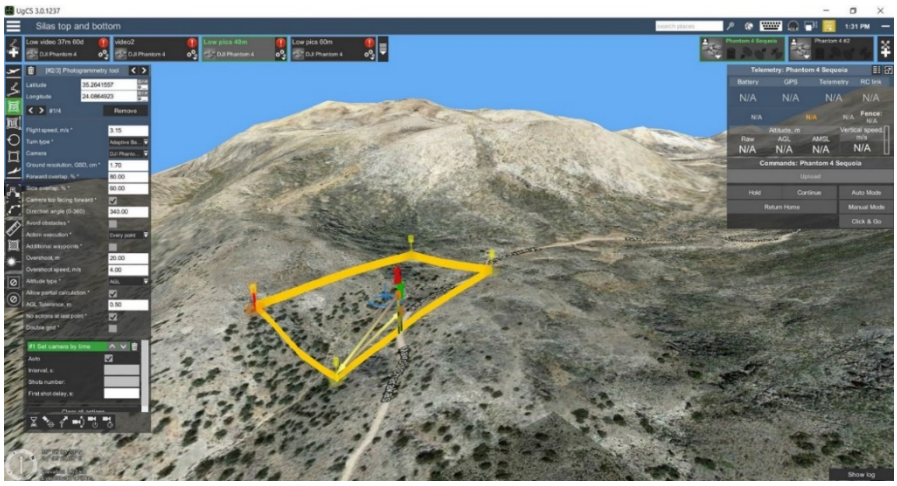


Figure 2.4. Telemetry window with navigation features

Figure 2.5: Interface of the UgCS software showing flight mission features. It enables terrain following mode flight mission planning

A total of five transects, approximately 0.6km by 0.2km were flown, and imagery was acquired twice for each transect at an interval of between 20-40minutes apart to generate 2 pairs per transect. 2 UAVs were used in taking both still and motion images. The average flying speed was 3.1m/s at a height of 40m. For each repeat mission per transect same flight plan used for the first image was used to reduce the effects of image to image registration. To improve image quality, images were not taken continuously during flight mode, but the flight speed stopped at each waypoint before taking an image. Each transect mission took a day. Half a day was used to acquire the images while the remaining time was used to manually review the

images after a flight to assess the quality of both still and motion images and whether repeat flight mission is necessary. Examples of raw images taken during the mission are shown in Figure 2.6.

Figure 2.6. Examples of raw UAV images taken. The image (right) shows the presence of animals on the ground.

2.2.3. Motion Images

Motion images were also taken using the same flight plan per transect. The videos were taken in between the transect pairs immediately after the first flight mission for a transect. The main reason was to capture as much as possible the features on the ground before significant animal movement occurs. A total of 5 videos were taken one per transect at the height of 40m.

2.2.4. Ground-based observations

Ground-based observations included counting of animals on the ground for each transect during the flight operation; and observing changes in weather such as cloud cover, precipitation and wind speed. During flight planning, care was taken to survey areas that had clear visibility from the home point of the Unmanned Aerial System (UAS). This enabled counting of animals within the selected areas. Most animals browse or graze as a group and the bells around their neck signal their presence within an area. Binoculars were used to focus on these areas where bells could be heard, and the animals counted. This was corroborated with video identification. Knowledge of the specific location where the animals were cited was therefore important for accuracy purposes when comparing with the manual counts from motion images. The aim was not to produce exact count but provide an estimate that can be used for subsequent analysis.

2.3. Data Processing

The main processing of the raw UAV images was carried out using Pix4D software. STAX and ERDAS Imagine software were also used in the processing stage. Preprocessing entailed generation of Point clouds, RGB mosaic images, DTM, DSM as well as motion Images. The RGB images were transformed using Independent Component Analysis and the nDSM generated was from DTM and DSM. Motion image was georeferenced using flight records. For this study, all data was projected to WGS 84 UTM Zone 35N.

The workflow for UAS data processing and analysis is presented in Figure 2.7. It also illustrates the image fusion process and change detection using Image differencing as well as accuracy assessment.

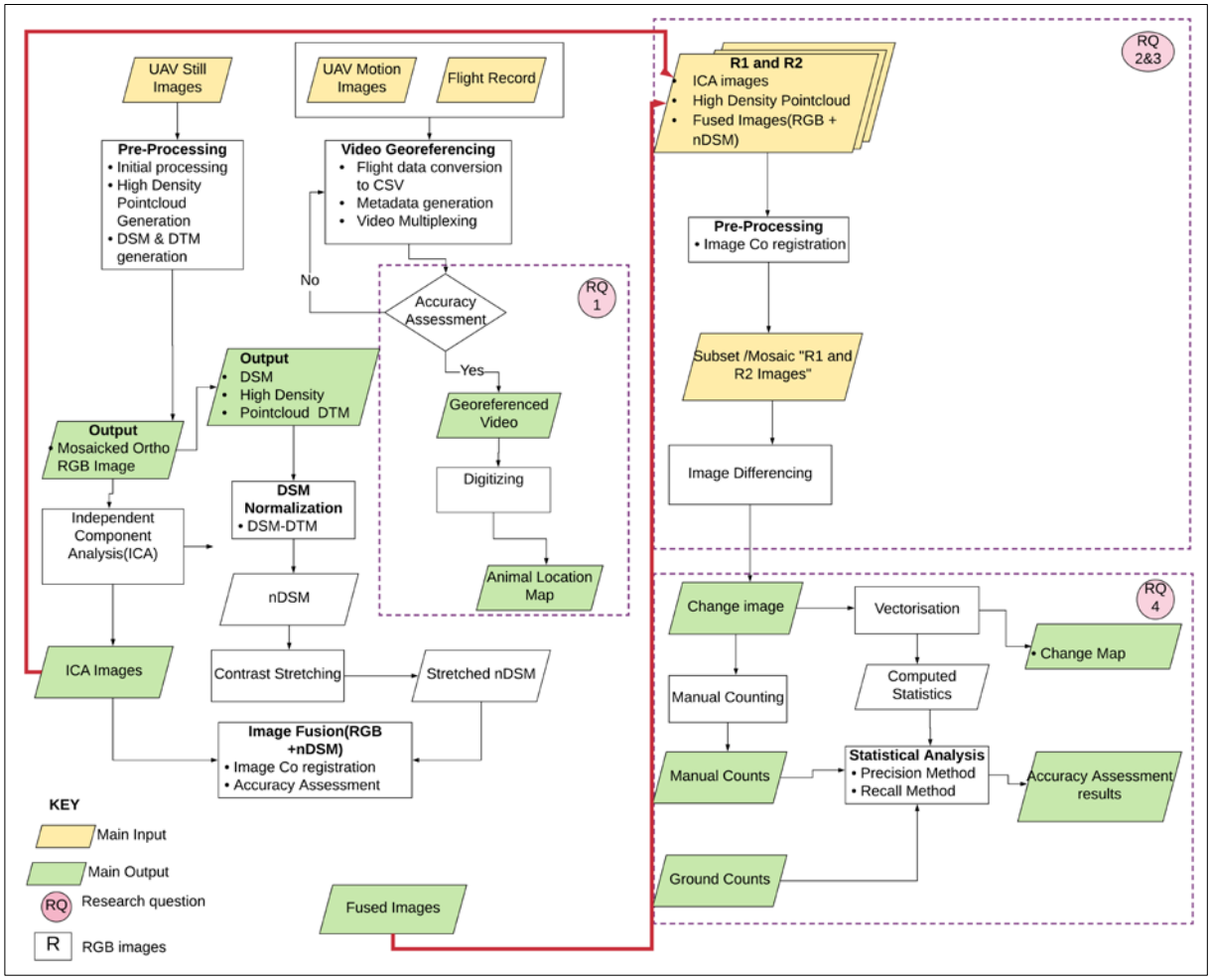


Figure 2.7. UAS data processing and analysis workflow

UAV images like satellite and aerial imagery are also affected by atmospheric effects. Before change detection, and depending on the analysis process to be done, it is essential that atmospheric, radiometric and alignment is performed. The following section describes the preprocessing steps; namely atmospheric correction, geometric correction; image matching and registration.

2.3.1. Image Correction

The main atmospheric correction entails correction of scene illumination, color correction and effect of strong sun intensity as well as cloud cover effect. Geometric correction can be either parametric or non-parametric. Since we did not use ground control points(GCPs) during flight mission, the parametric approach was employed which uses exterior orientation parameters provided by the Initial Navigation System (INS) (Shi et al. 2011). For image matching, bundle block adjustment (BBA) method which uses corresponding key points and image coordinates to match overlapping images together before mosaicking was applied. Structure from Motion was used to estimate the 3-D structure of a scene from a set of 2-D images. All the processes were done within Pix4D software.

Pix4D is widely used in the processing of UAV images. It processes images taken from any angle from a manned or unmanned platform and automatically calibrates to compensate for illumination and color changes. Also, it radiometrically corrects for image reflectance. The software also runs an Automatic Aerial Triangulation (AAT) and Bundle Block Adjustment (BBA) automatically using either parametric or non-parametric approach and the result is a radiometrically and geometrically corrected images in 2D, 2.5

and 3D images such as Optical images, DSM, DTM, point cloud and 3D mesh in tile and mosaic form (Figure 2.8. and Figure 2.9.)

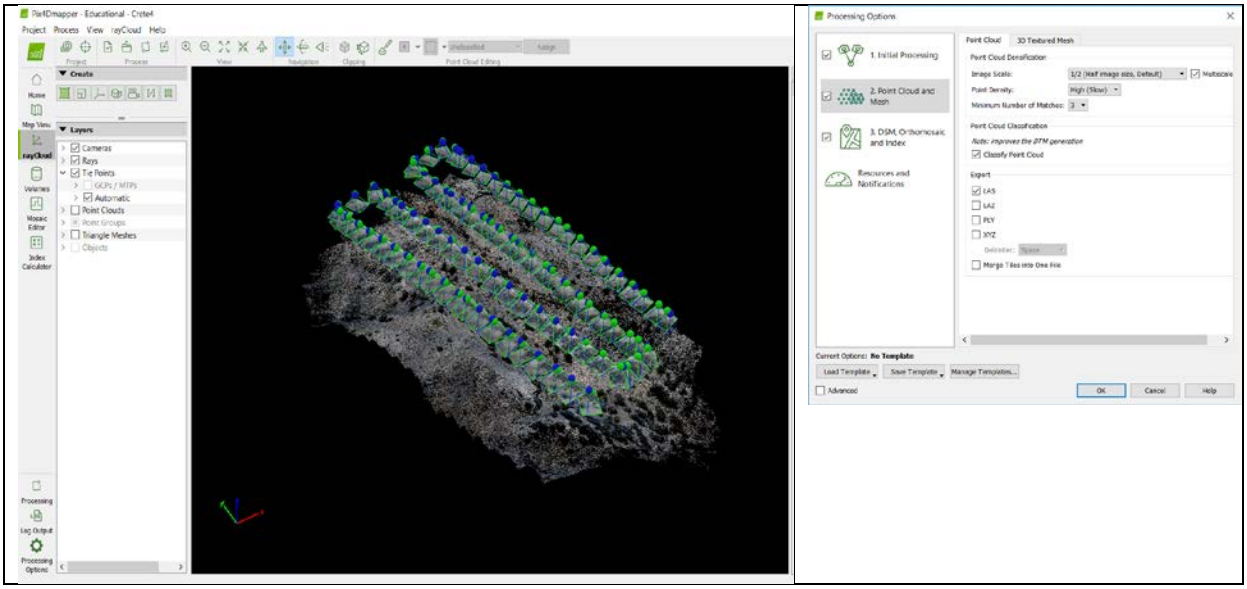


Figure 2.8. Pix4D image processing.

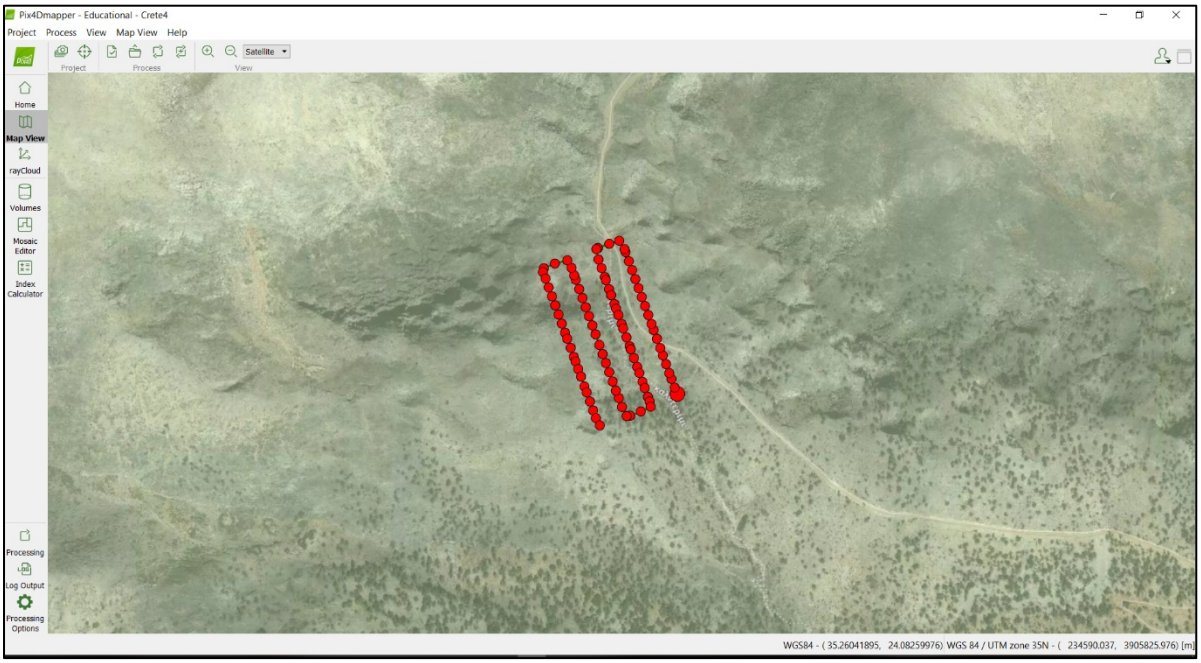


Figure 2.9: Flight path and waypoints. Pix4D has background images of an area enabling visual presentation of an area covered by the flight mission.

2.3.2. Image alignment

The section entails co registering of two pairs of images of a transect through geometric transformation. The positions of pixels which are defined by scale, rotation and translation are modified using mapping functions and these mapping functions pairs are used to relate corresponding pixels in the two images.
According to Hollings et al., (2018) UAV images have very high resolution, and most of the images cannot be effectively georeferenced and co-registered for change detection. The submeter accuracy does not provide enough georeferencing information to match two images of the same place accurately. This requires GCPs which is time-consuming and in situations where the GCPs were not taken, the accuracy of

the detection process is limited. However, some researches have been done, and new techniques developed to automate the co-registration of image blocks without using the GCPs. For image alignment, STAX software from PCI Geomatics was used. The software automates the alignment process and uses a designated image between the pairs as the reference image thereby eliminate the use of ground control points (GCPs) (Figure 2.10.) Alternatively, a highly accurate reference image of the same resolution as the image pairs can be used to align.

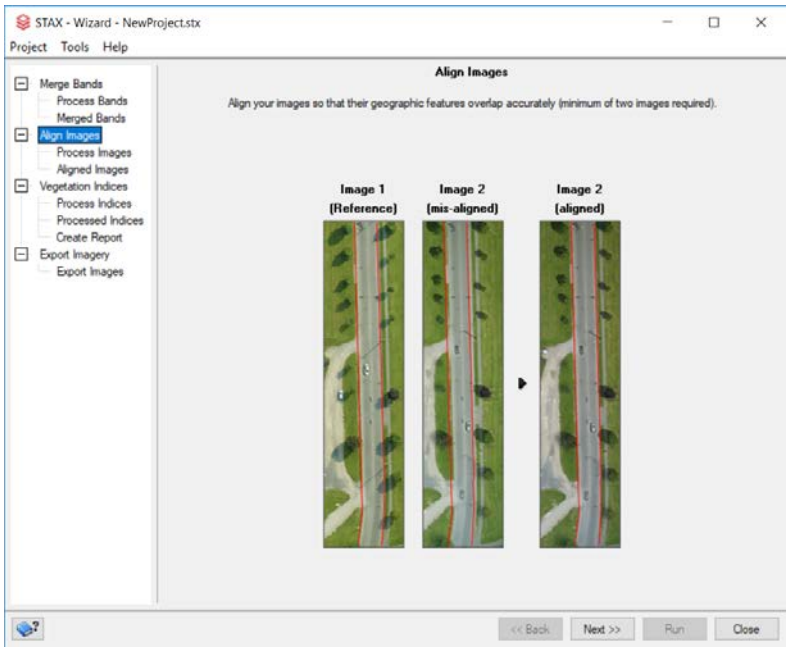


Figure 2.10: Image alignment process window showing an example of the alignment process and the outcome. STAX software specifically processes UAV images.

Two image pairs of a transect were loaded into the software as image 1(R1) and image 2 (R2). R1 image was used as the reference image. After alignment, the results were exported and saved in tiff format.

2.3.3. Motion Images Georeferencing

Videos were also generated per transect for subsequent analysis. To perform video georeferencing, metadata information such as the xyz position of the drone, attitude, tilt, camera angle, and camera characteristics are required in line with the Motion Imagery Standards Board (MISB) standard (MISB 2017). These standards were developed for capturing videos with metadata that enables georeferencing. The georeferencing was done using Full Motion Video in ArcGIS software. Details of the process are explained in the subsequent method section.

A summary of the data and software used for the study is presented in Table 2.2.

Table 2.2. The table shows a summary of data collected for the study

Data	Quantity	Resolution
Mosaicked UAV images	12	3cm
Motion Images	6	3cm
Ground Counts (2 transects)	21 &18	

Software	Purpose
ArcGIS 10.6.1	Data analysis and presentation
Full Motion Video (ArcGIS Extension)	Motion Image georeferencing
ERDAS Imagine 2018	Data analysis
Universal Ground Controls Station (UgCS)	Flight mission planning
Pix4D version 4.3.27	UAV Data Processing
STAX	UAV Image alignment
Excel 2018	Data analysis

2.4. Data Analysis

Based on the research questions of this study, the methods are divided into 4 main parts. The first part focuses on assessing the spatial distribution and locations of animals using motion images and its measure of performance. The second part focuses on change detection using image differencing by applying Independent Component Analysis, and the third part looks at change detection of fused images (RGB image and nDSM). The last part assesses the feasibility of the image differencing method based on the second and third questions to determine whether the accuracy rates are high enough (≥50%) for wider application in rangeland management.

2.4.1. Using motion images for spatial location of animals

Direct Georeferencing using Inertial Navigation Systems and Global Navigation Satellite System (INS/GNSS) sensor is possible with the aid of flight attitude and video frame parameters to compute the relationship between the map coordinate system and camera coordinate system for each frame (Eugster & Nebiker 2008). Georeferencing the motion images was done using ArcGIS full motion video tool based on the MISB standards. The main idea was to use the georeferenced image to digitize animals on the video to portray patterns of location and abundance. ArcGIS Full Motion Video (FMV) allows running of georeferenced videos on the map view. A smaller window displays the video as it runs while the map view displays the footprint(rectangle) of the video on the map view as the video is run. A short line arrow shows the flight path of the drone. Any process such a digitizing feature on the video is also shown on the map view and vice versa. The workflow of the process is presented in Figure 2.10.

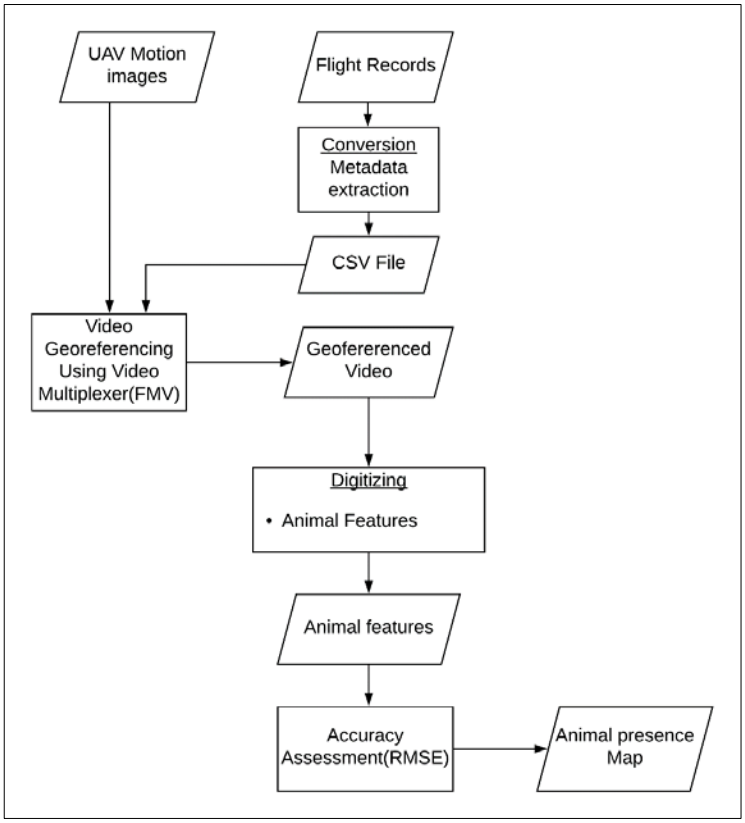


Figure 2.11. Video georeferencing process.

Video Georeferencing process

Metadata for the videos in MISB format were used for georeferencing. Most UAVs taking videos are MISB non-compliant therefore the motion images are not georeferenced, but the information on the flight log is used to create metadata which is then integrated with the video. Metadata DJI flight logs are encoded therefore cannot be directly used until they are decoded using flight log viewer and converted to readable CSV files. https://www.phantomhelp.com/logviewer/Upload/ website was used to load the flight logs and generate a CSV file for further processing. The verbose CSV contains all the information and details such as the xyz position of the UAV, attitude, tilt, camera angle, and camera characteristics as shown in Figure 2.11.

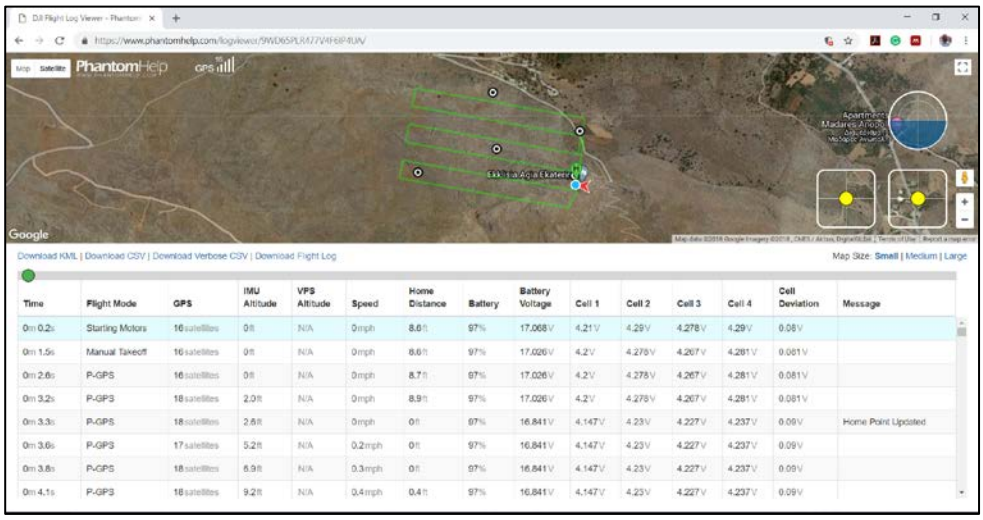


Figure 2.12. Flight log viewer and processing interface

For georeferencing using FMV, the following flight log information are required; UNIX Time Stamp; Sensor Latitude; Sensor Longitude; Altitude; Heading; Pitch; Roll; Horizontal Field of View (FOV); Vertical Field of View (FOV); Sensor Relative Azimuth; Sensor Relative Elevation and Sensor Relative Roll.

These 11 fields are important in creating a video frame footprint which is also the frame outline on the map. UNIX time Stamp also known as the Coordinated Universal Time is represented as the number of microseconds elapsed since midnight of January 1970. For sensor position calculation; sensor latitude, longitude, and altitude are required. The orientation of the camera were provided by attitude information which includes platform heading, pitch and roll. Frame lines were defined using the horizontal and Vertical field of View. To define the video corners that will be projected on a map view, relative sensor orientations which is described by the relative azimuth, relative elevation, and relative roll were used.
The Verbose CSV file generated from the DJI flight log viewer was cleaned to remove fields that were not required. See Appendix A.

ArcGIS Full-motion Video Multiplexer was used to generate a video encoded with the metadata. The resultant georeferenced output was displayed using FMV add-in tool (Figure 2.13). To verify the accuracy of the result, an orthorectified still image of the transect was also displayed on the map view, and the video was set to run to check whether the footprints fall within exact areas on the orthoimage. Points were marked and checked if they correspond to the still image on the map view.

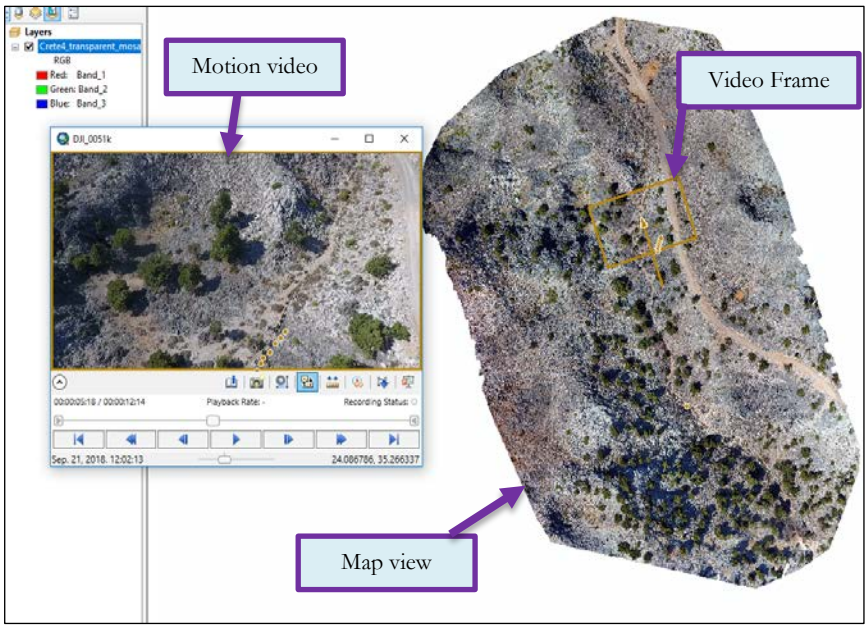


Figure 2.13. Video frame and map view. Direct digitizing process

Direct Digitizing

The area under study has a low contrast, therefore, differentiating animals from background was challenging. Animals on the video were identified due to their movement, and points representing animals on the video were marked and automatically shown on the map view. The points were later saved in a file geodatabase as vector files.

Accuracy Assessment

To test for the accuracy of the georeferenced video. Known control points representing observable features that were not animals on the map view were identified and digitized on still image on the map view. Corresponding points for the control points were also digitized on the motion video and compared. Discrepancies between the points in terms of distance were calculated and recorded. A Root Mean Square was calculated between observed values on the still image and acquired values on the motion image. The equation used to calculate the RMSE is indicated below.

$$m_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \delta x_i^2}$$

Equation 2.1. RMSE x direction

$$m_y = \sqrt{\frac{1}{n}\sum_{j=1}^{n} \delta y_i^2}$$

Equation 2.2. RMSE y-direction

Where m_x represents RMSE in the x direction and m_y represents RMSE in the y direction.
The overall RMSE was calculated using the equation below.

$$Overall\ RMSE = \sqrt{m_x^2 + m_y^2}$$

Equation 2.3. Overall RMSE

A low RMSE indicates better results.

2.4.2. Image differencing using Independent Component Analysis (ICA) images.

This section focuses on image differencing on transformed images using Independent Component Analysis and its feasibility to provide an overall detection rate comparable to a ground-based survey in the detection of animals. Conventional methods for linear transformation include principal component analysis, factor analysis, and projection pursuit. Independent Component Analysis also known as blind source separation technique is a new method that recovers linear transformations of non-gaussian data so that they are statistically independent (Hyvärinen & Oja 2000; Mirvakhabova et al. 2018) ICA is, therefore, a high order feature extraction technique and offered a better way of reducing spatial dimensionality.

While PCA decorrelates the transformed variables, ICA tries to make the variables as statistically independent as possible by minimizing their redundancy or common information (Mirvakhabova et al. 2018) . Lee & Lewicki (2002) reported in their research on using ICA models in unsupervised image classification that ICA models provided more information on image features than other gaussian models. Jiménez-Hernández (2010) also used subtraction approach in object detection based on Independent Component Analysis. ICA, therefore, is ideal in this study because of better pattern recognition and consequently useful in feature extraction and classification.

The analysis was done on RGB images and point cloud separately as well as using a combination of RGB and normalized DSM. The workflow is illustrated in Figure 2.14.

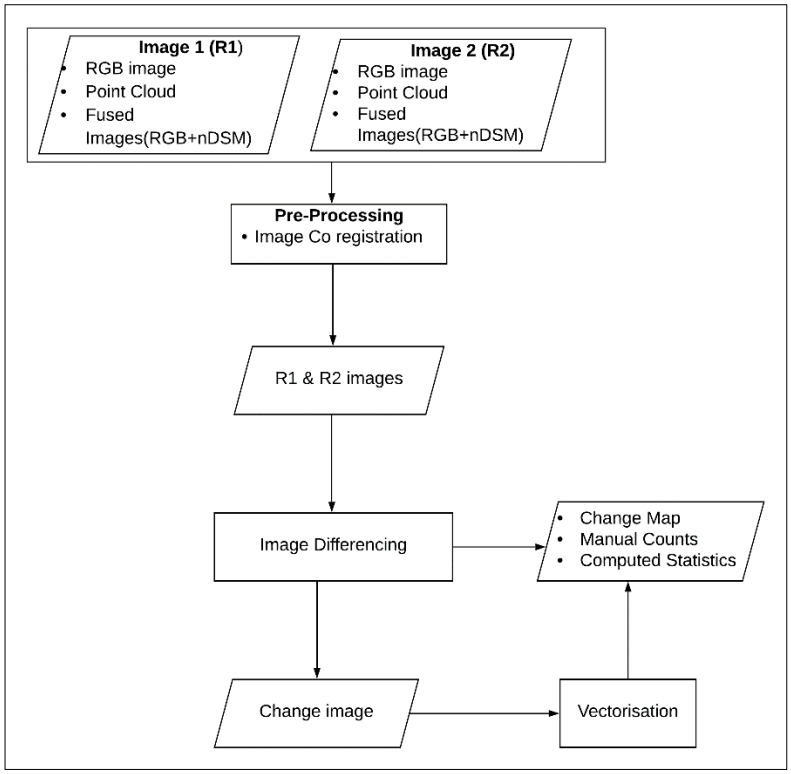


Figure 2.14. Image differencing for an optical image pair

A pair of 6 images representing the transects were selected for data analysis. Since not all images had animals, only 2 transects with 2 image pairs having animals were selected for further analysis because they had the presence of animals based on ground counts. The first step was to run ICA on each image to generate original and decorrelated features. One image within a pair was defined as the R1 image and image taken after the short interval was defined as the R2 image. A subset was created for each pair focusing on areas that animals were identified on the ground. The aim was to focus on a small area in image differencing using ICA and later a large mosaic area using the same method and compare the results. A total of 5 sites were created as per Figure 2.15. Both for R1 and R2 image. see Appendix B. for the other 4 sites.

RGB Image 1(R1) **RGB Image 2 (R2)**

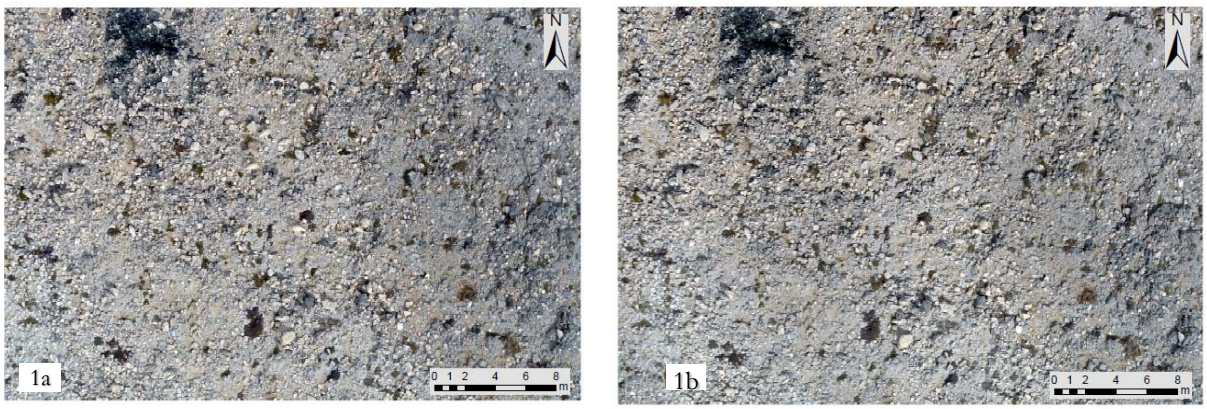


Figure 2.15. Image from site 1 selected for the study.

A differenced image result was obtained by subtracting the ICA image R1 from the ICA image R2 for each subset of a transect using the minus tool in ArcGIS. The difference was dictated by pixel to pixel relationship, and high absolute differences in image values represented animals detected while low differences represented no or minimal change. The assumption was that high pixel-pixel difference values will represent changes on the R1 image which is our reference image while low pixel-pixel difference values will represent the differences in the R2 image. High pixel-pixel differences were identified as animals and compared to ground and manual counts based on observed locations of the animals.

2.4.3. Thresholding and Vectorization

For the purposes of this study, we defined thresholding, vectorization and subsequent identification of the polygons representing animals as semi-automatic processes. To semi-automate the detection process, pixels having high intensity values and falling within locations where animals were spotted and positioned were converted to polygons. First, standard deviations of the differenced raster layer output were recorded and since we are interested in high contrast difference, positive pixel intensity values above the standard deviation were assumed to represent change. A threshold based on the standard deviation value was set to isolate pixel values below the standard deviation. Standard deviation measure how spread out the variation is from the mean. A low standard deviation indicates that the values are close to the mean, while high standard deviation indicates high spread out (Kumar & Gupta 2012).

The expectation was, short time interval image differencing will generate minimal differences except for changes in locations where animal movement occurred. Most values would, therefore, fall closer to the mean, and any value above the standard deviation is treated as an anomaly which closely relates to the animal movement. Only positive changes above the standard deviation were factored. Thresholding was done using a spatial model in ERDAS Imagine (Figure 2.16.). The result which is continuous was converted to a thematic layer, and a majority filter of 3x3 was run to eliminate noise.

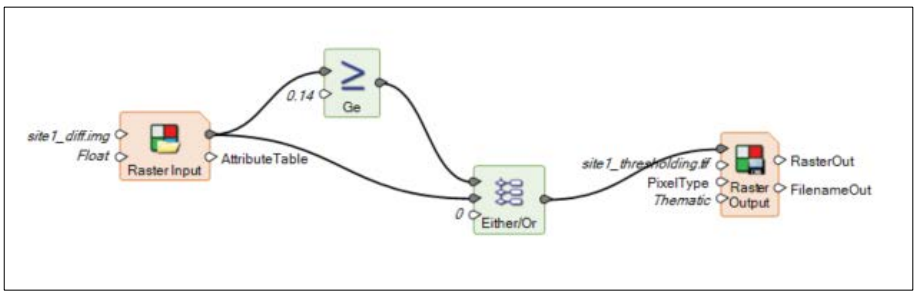


Figure 2.16: Raster layer thresholding spatial model. The standard deviation value is set using the ≥ operator.

The thematic layers were vectorized using ERDAS Imagine. Gaps and silvers were removed in the vector layer using the eliminate the polygon part tool in ArcGIS. Thereafter the dissolve tool was used to merge and remove holes. Second thresholding was done using ArcGIS software based on an approximate physical size of the medium-sized animals. Approximate size of between 0.1m² and 0.3m² was set for the animals based on the measurements done on known animals on the high-resolution images using the measurement tool in ArcGIS although this had the potential of eliminating very small animals. Polygons outside the range were removed, and the remaining polygons within the range were considered potential animals. These polygons were counted and compared with manual and ground-based counts to generate true positives, false positives and negatives.

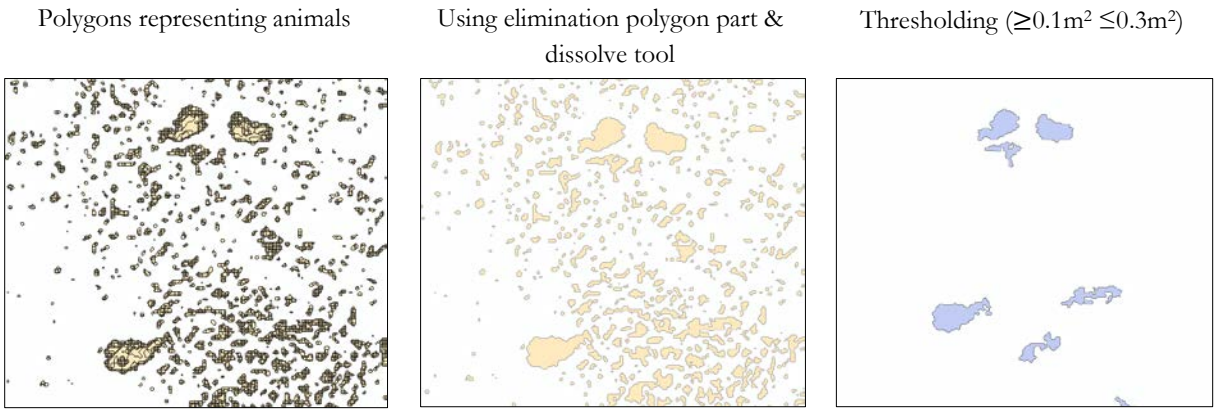


Figure 2.17: vector cleaning process. The thresholding is based on a measured value representing the size of the animals on the ground which was between 0.1m² -0.3m². Values outside this bracket were removed from the vector layer.

2.4.4. Using point cloud to point cloud comparison (Point cloud approach)

To further assess whether products generated from point clouds can be used to improve detectability, were compared point clouds of transect image pairs representing R1 and R2 image for the selected sites that had an animal presence. According to Girardeau-Montaut et al. (2005), in their research on change detection on points cloud data acquired with a ground laser scanner, a direct point cloud to point cloud comparison using two datasets acquired of the same area is possible and these results can be used as a preview process or onsite verification before other processes are applied. The aim was to detect changes from high-density point clouds to determine whether their products can be fused with RGB images for the change detection process.

Point cloud change detection was done by comparing two point cloud datasets of the same area taken at different intervals. This technique entailed computing point cloud to point cloud distances and one point cloud layer acting as the reference. The general assumption is we expect the point to point distances to be very narrow for the same pairs of point cloud images taken the same day for a given transect but due to animal movement, anomalies are expected in areas where animals moved, and these discrepancies are assumed as animal locations. This process was done using CloudCompare, an open source software for processing and analyzing point cloud data (Figure 2.18.). It offers various algorithms that are useful in point cloud image registration and alignment. It also offers distance computation either point cloud to point cloud or point cloud to mesh. Lastly, it provides statistical computation and segmentation as well as visualization.

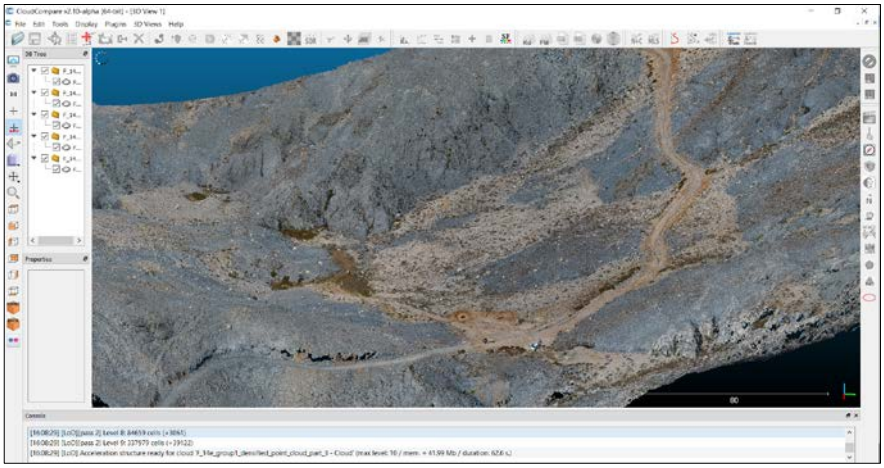


Figure 2.18: CloudCompare interface with a point cloud image of the study area.

First, the two R1 and R2 point clouds datasets representing a transect were aligned, and RMSE calculated to assess the accuracy of the alignment using cloud compare. Low RMSE meant the aligned results are good. Only pairs of transects where animals were observed were used for comparison. A point to point comparison was done on pairs of point cloud images representing a transect. The image and statistical results generated were analyzed based on distance from the mean. The standard deviation generated was used to set maximum threshold values. Point to point distances outside the set threshold was interpreted as a change. Points distances less than the standard deviation and close to the mean were deemed as no change. Girardeau-Montaut et al. (2005) further indicated that it is possible to set threshold values using the standard deviation and visualize as well as process only points whose distance to nearest neighbor in the other point cloud image is less or more than the standard deviation value.

2.4.5. **Using** fused images (RGB + nDSM)

To further Image fusion has been used in many remote sensing applications. The aim of fusion is typically to improve the accuracy of results by integrating additional information from various sensor information into one image (Prošek & Šímová 2019). This is particularly important when we want to compensate the weakness of one data source against the other and vice versa.

The method flow included the use of normalized DSM generated by subtracting DTM from DSM; aligning; calculating RMSE and Integrating nDSM with ICA image generated for R1and R2 images for each selected subset with animals. Lastly image differencing of the resultant images. The assumption was that integration of elevation information from nDSM with pixel intensity information from UAV RGB image provides additional information for improving detectability. Each step is defined as follows based on the workflow in Figure 2.19.

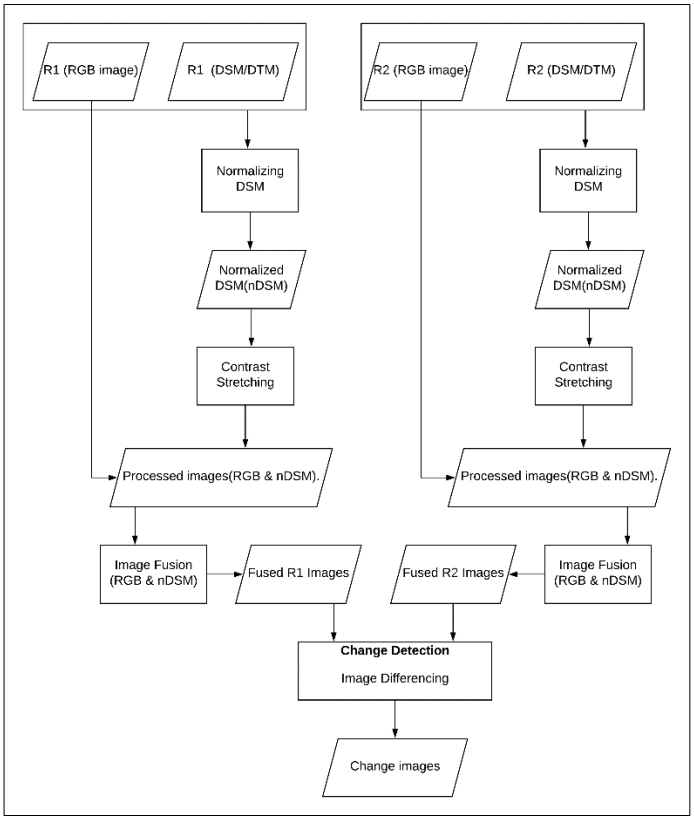


Figure 2.19. Flowchart of the image fusion and change detection process.

2.4.6. Normalized DSM (nDSM) generation

The idea was to extract heights above the ground that can be used to detect animals based on changes in height between images. Experiments were done with DSM and the results were not promising therefore nDSM was chosen. nDSM represent a micro topography with information on heights of objects. The following input was used for this section. ICA images of sites having animals. DSM and DTM corresponding to the ICA images of selected sites containing animals. To calculate nDSM, DTM for each site were subtracted from the DSM for each corresponding site (Figure 2.20). The nDSM for R1 and R2 image were aligned using STAX software to their corresponding ICA images.

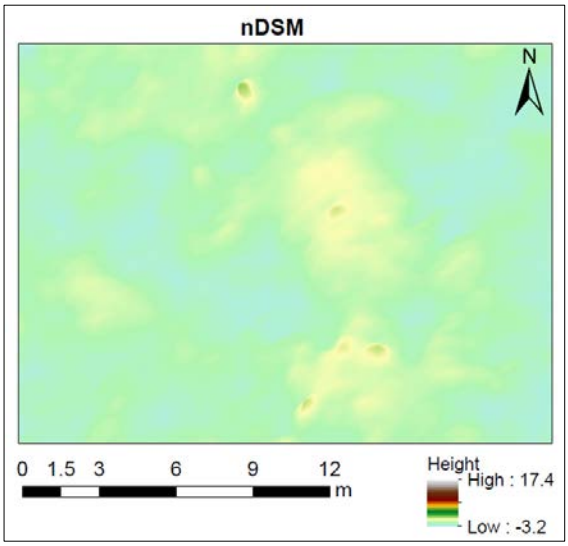


Figure 2.20. nDSM representing site 2.

Image Fusion

ICA images generated from RGB optical images had a scale range of 8 bit (0-256). It is imperative that before the fusion of these datasets, both RGB and elevation data should be on the same scale range because the pixel values in both raster layers represent different information. nDSM raster image has elevation values while the ICA image has pixel intensity values. The nDSM images were therefore scaled to a similar range as the ICA images. This contrast stretching was done using a spatial model created in ERDAS imagine as shown in Figure 2.21.

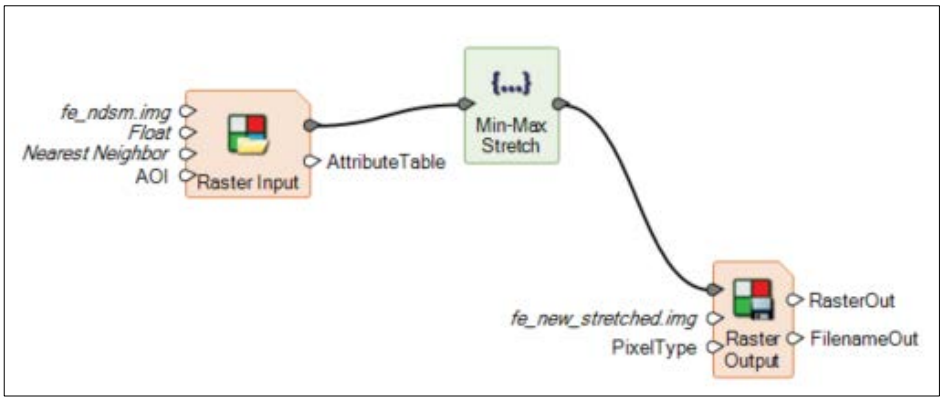


Figure 2.21. A spatial model for contrast stretching.

The ICA images were then integrated using ERDAS Imagine to their corresponding nDSM by stacking the two layers. The output integrated images pairs for each transect were aligned using STAX software.

Image difference of the fused images for R1 and R2 were made and detected animals were compared to manual counts and ground counts. Vectorization and reclassifying were done the same as the previous operation with RGB images. True positives, false positives and negatives were calculated and the results from the methods used compared and analyzed.

2.4.7. Accuracy Assessment

In order to assess the accuracy of the methodology used, ground and manual counts were compared with results from the change detection method used on the RGB and fused image. Figure 2.22. presents the workflow that was used.

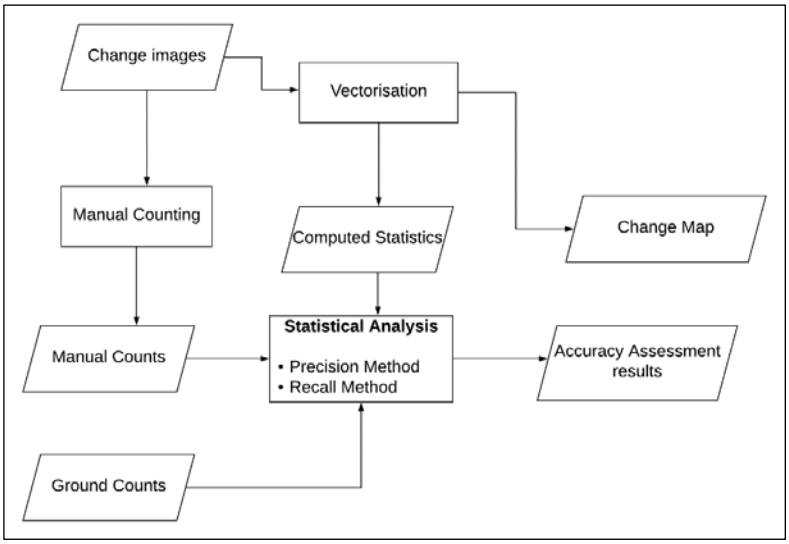


Figure 2.22. Accuracy assessment process.

The accuracy assessment entailed computation of True Positives (TP), False Positives (FP) and False Negatives (FN). True Positives, in this case, were matched animals based on the manual and ground count with the results from the vectorized image. False Positives were polygons representing animals that were determined on the output image but are missing in the ground and manual data. False Negatives are animal features present on the ground and manual data but cannot be identified on the output image (Sarp et al. 2014).

2.4.8. Precision, Recall and F-measure

Precision and recall index was used to estimate the accuracy of the methodology. Precision generally is concerned with the exactness or closeness of measurements and it is independent of accuracy. Recall also known as specificity, on the other hand, is more concerned with the completeness or level of details of a given measure(Khan et al. 2018). The F score is a harmonic mean of precision and recall. The formula for calculating precision, recall and F measure are given below;

$$Precision = \frac{TP}{FP + TP}$$

Equation 2.4. Formula for calculating precision

$$Recall = \frac{TP}{FN + TP}$$

Equation 2.5. Formula for calculating recall

$$F\ score = 2\frac{P * R}{P + R}$$

Equation 2.6. the formula for calculating F score

These three methods were chosen because they generate objective assessments when looking at both the qualitative and quantitative aspect of the method used. Qualitative because it focusses on the trueness of the result and quantitative because it focusses on the number that has been detected (Powers & Ailab 2011). Precision, therefore, is qualitative while recall is quantitative.

The general assumption is that high precision implies, the method produces more positive results than negative while high recall implies, most of the positive results are identifiable. Therefore, high precision and recall is a good measure. F measure was an equal-weighted score used to have a balance between precision and recall.

3. RESULTS

This section presents different outputs generated based on the research methodology and includes video georeferencing results as well as validation; change detection using image differencing on UAS images, point cloud and a fusion of optical and nDSM. Lastly, the section presents the accuracy assessment of the manual and semi-automatic method used in detection and counting and compares RGB and fusion approach.

3.1. Animal presence data collection using motion images

Video multiplexer was used to georeferenced motion images by integrating metadata from flight records and a motion image of a transect to generate corner coordinates and frame lines. Figure 3.1: FMV.shows the result of the process.

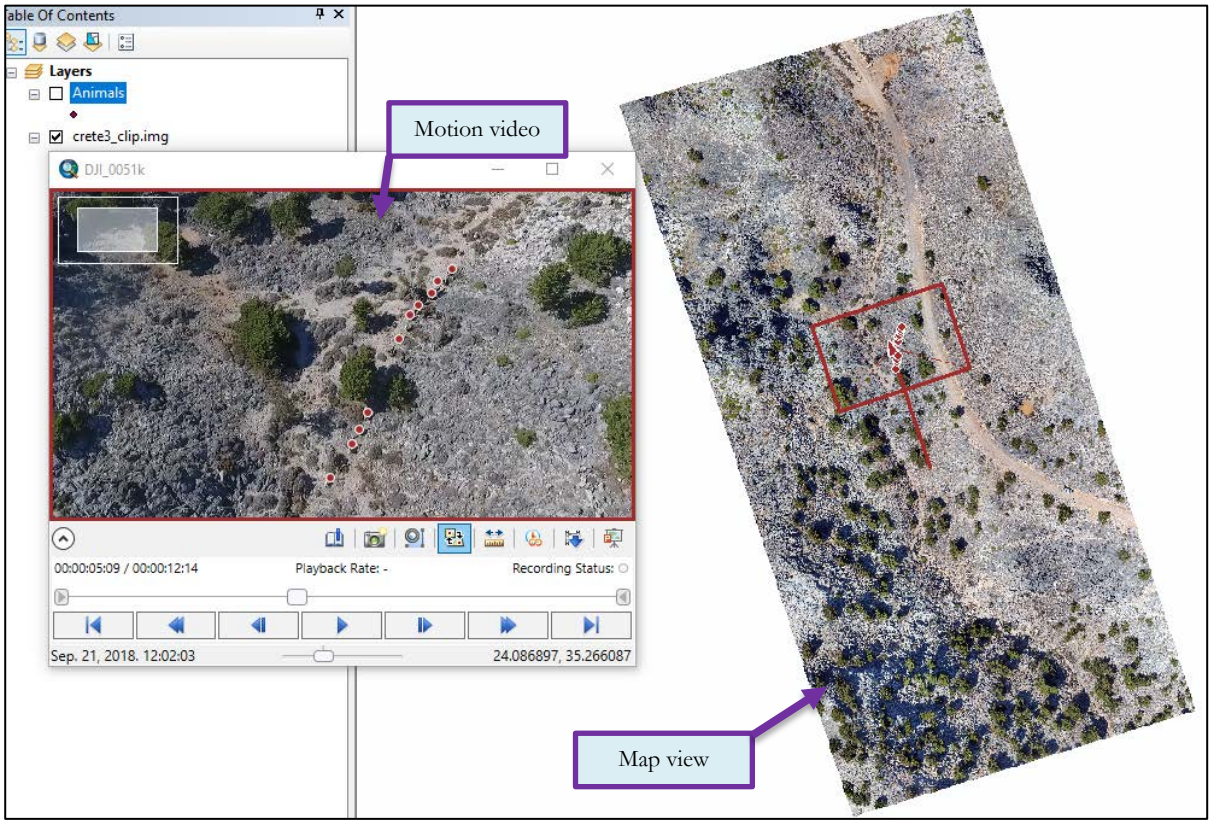


Figure 3.1: FMV window and map view showing the digitization process of the animals which are simultaneously updated on the map view as represented by the red box on both the video multiplexer window and map view.

3.1.1. Digitizing

A total of 25 animals were digitized from the video. Although the animals were moving during the digitizing process, the frame was paused, and digitizing done. The result included their distribution on the map view as presented in Figure 3.2. The results generally presented an estimated location of the animal directly georeferenced from the video frame.

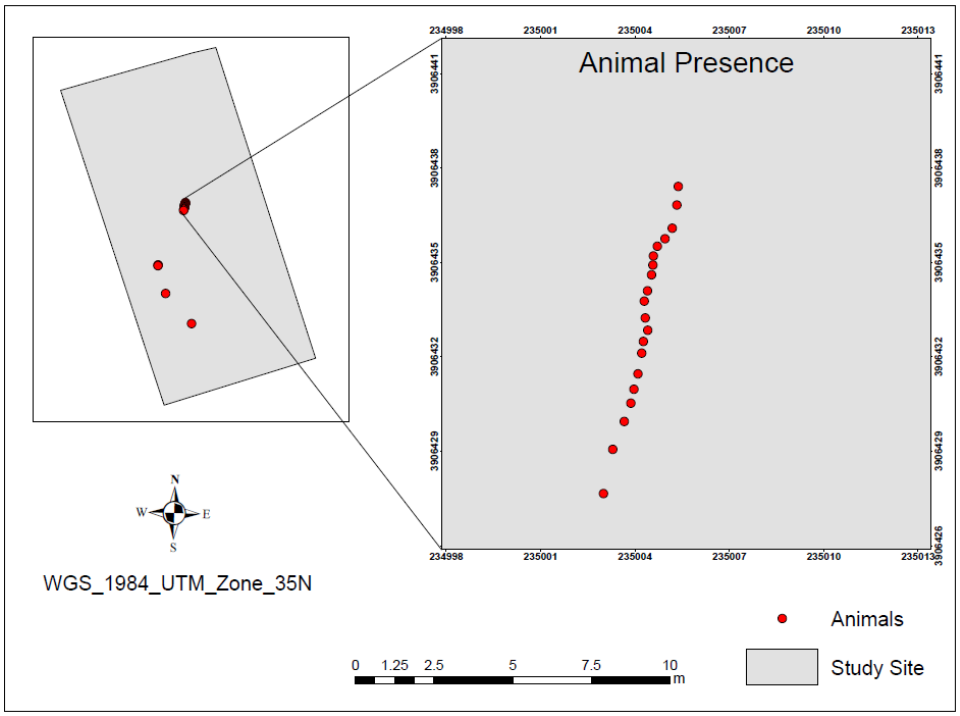


Figure 3.2: Map showing animal presence digitized from the motion image.

3.1.2. Accuracy Assessment (Validation)

Root mean square error was used to validate the results to check on the precision and accuracy. It was used to evaluate the deviation of the derived motion image points from the measured Ground Control Points using the formula indicated under the methodology section. Location data was subsequently generated from the motion and map view. Figure 3.3. shows the reference points and motion image points that were used including their distribution. A total of 20 corresponding validation samples of known and visible features were manually digitized both on the map view and corresponding motion image.

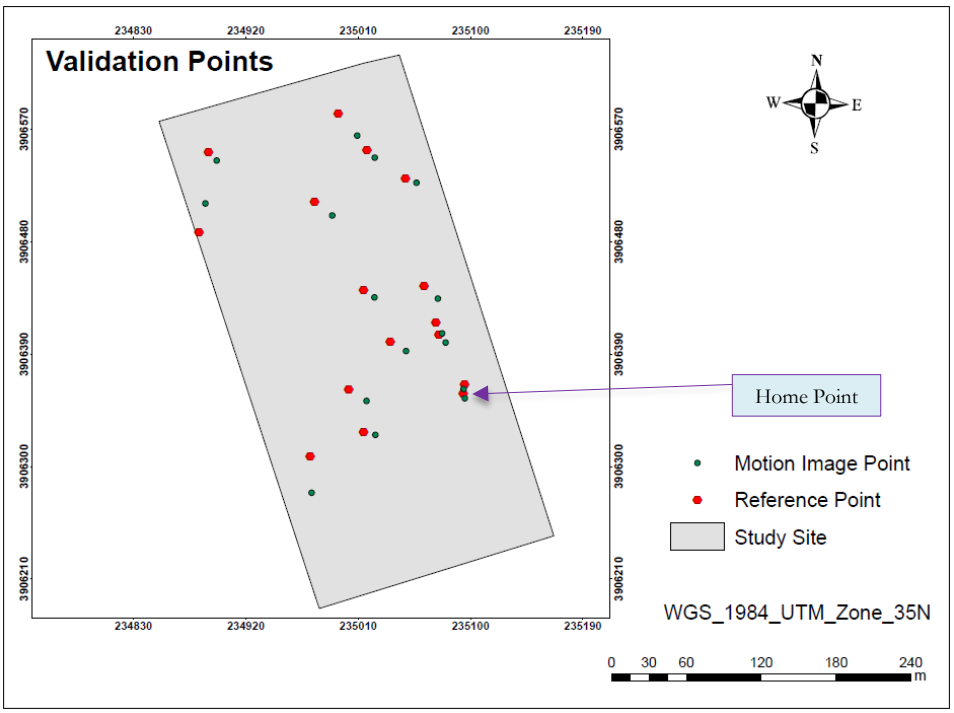


Figure 3.3: Location and distribution of validation samples

An RMSE of 9m and 12m were generated for y and x direction. The overall RMSE was 15.12. Areas closer to the home point had lesser differences than those in rugged and steep areas. It is worth noting that the motion image though georeferenced, it is not rectified; therefore this accounts for the large differences between derived and measured and consequently the large RMSE. The map view image where reference points were taken is orthorectified.

3.2. Image differencing using RGB image

For this study, the average size of site 1-4 was 0.1ha while site 5 was 9.8ha. ICA was applied to generate uncorrelated images shown in Figure 3.4 and 3.5. The images pairs were aligned using STAX software and Image differencing was subsequently done for the selected sites and results generated as shown in the figures below. High values indicate potential animals. Potential animals that were on the second image are shown as depressions with low values as shown in Figure 3.5. Animals that did not move were not detected in some sites.

RGB image *Image difference*

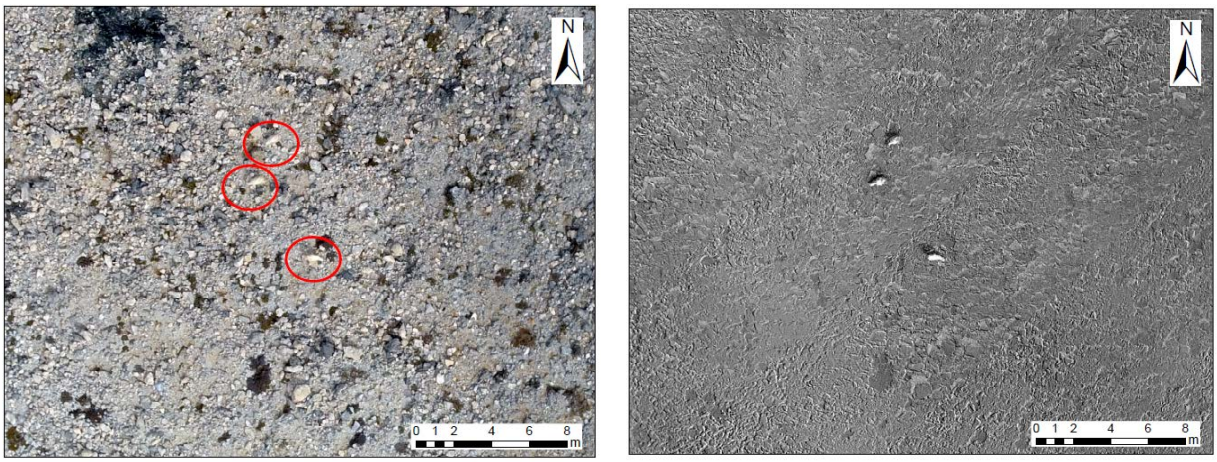


Figure 3.4. Site 1 image original (left) and difference image (right). Red shapes indicate animal locations

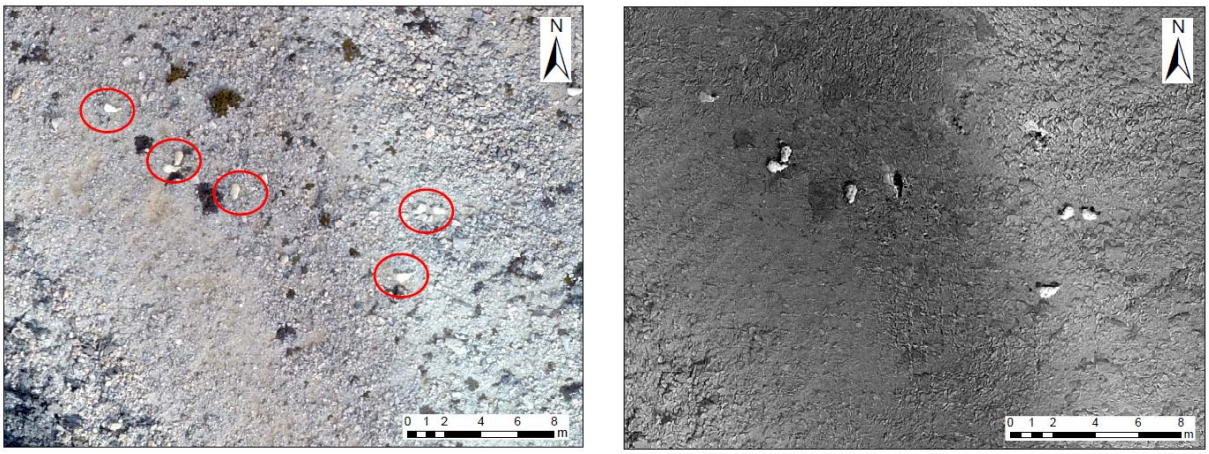


Figure 3.5. Site 2 image original (left) and difference image (right). Red shapes indicate animal locations

3.2.1. Manual and Ground Counts

In this study, manual counting of the detected animals was done using the count feature tool in ERDAS Imagine. This tool enables counting of features accurately, and referral can be made in case of omission. The results were compared with ground counts per selected site (Table 3.2). A total of 16 animals were manually counted on the differenced images. Ground counts representing the total number of known

animals was 18. This shows a difference of 2 animals between the manual and ground count which could be as a result of image preprocessing errors and some animals did not move.

Table 3.1. manual counts for selected sites

Site	Size(ha)	Manual Count	Ground Count
1	0.1	3	3
2	0.1	7	8
3	0.1	3	3
4	0.1	3	4
5	9	16	18

3.2.2. Semi-automated count

Semi-automatic counting was done by vectorizing the differenced images from each site. Although this method is an evaluation of the image differencing using thresholds to assess the changes representing animals, the name "semi-automated method" was used to differentiate it from manual counting of the animals on a differenced image. First, a standard deviation stretch using ERDAS imagine was applied, and the standard deviation and the mean for each site were recorded. Generally, in the selected sites, most values were within the standard deviation.

The standard deviation ranged between 0.1 and 0.7 as per the trend observed on Table 3.3. and most values fell around the mean which we can assume to represent no change.

Table 3.2. Threshold for TP of selected sites for RGB image method

Site	Mean	Standard Deviation (Threshold)
1	0.3	0.7
2	0.4	0.4
3	0.02	0.6
4	0	0.16
5	0.02	0.1

Thresholding was done based on the above standard deviation for each site to isolate the pixel values that do not represent change. Conversion to thematic layer, vectorization and subsequent cleaning of vector layer as well as further thresholding based on the animal size of $\geq 0.1m^2$ and $\leq 0.3m^2$ was applied to reduce FP. Figure 3.9 - 3.13 represents the results from the described process.

Site 1

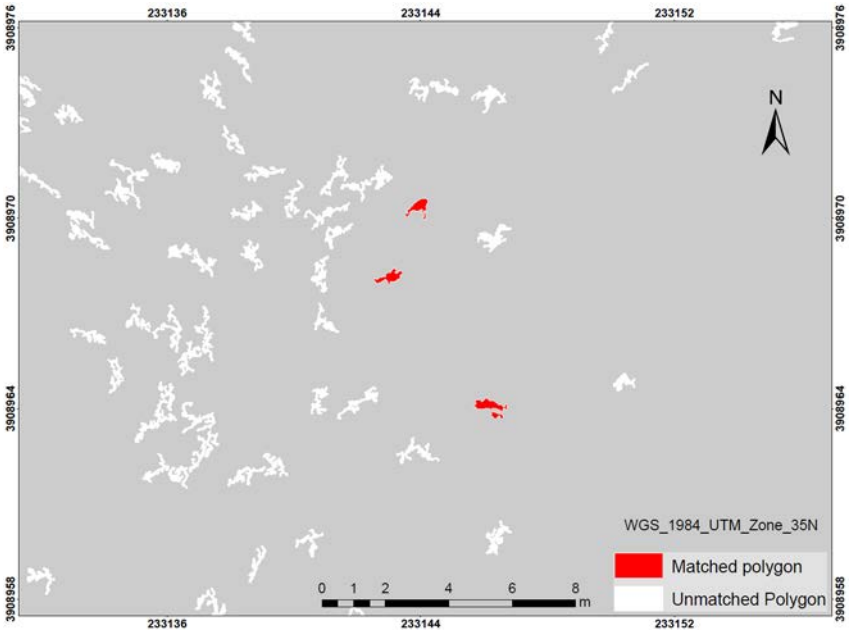


Figure 3.6. Map of identified animals (matched polygons) on site 1 using one standard deviation threshold (0.7) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m².

Site 2

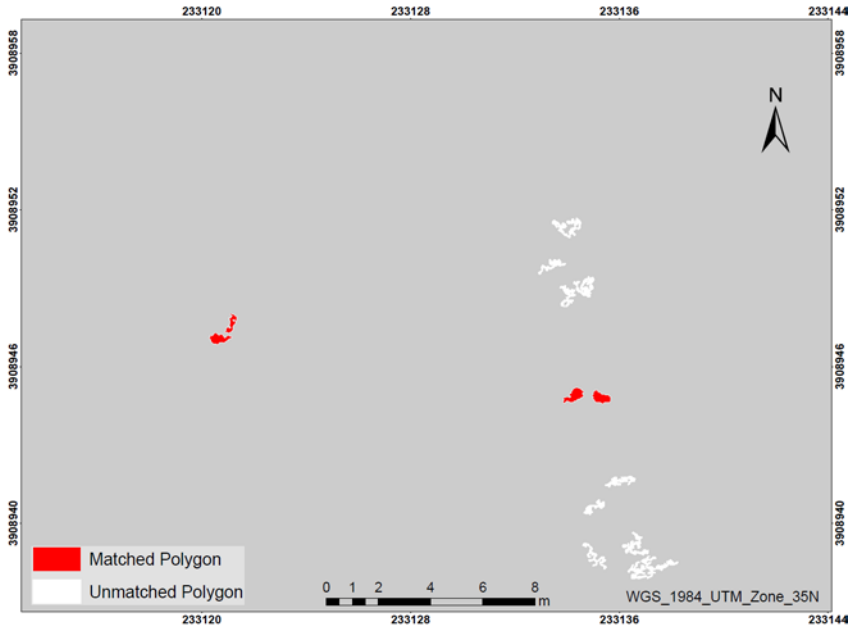


Figure 3.7. Map of identified animals (matched polygons) on site 2 using one standard deviation threshold (0.4) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m²

Site 3

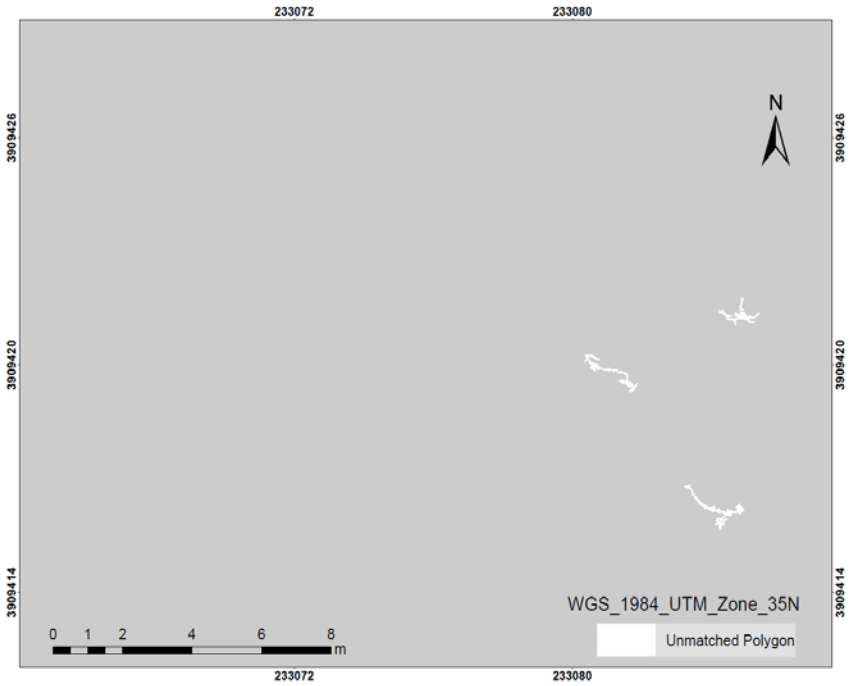


Figure 3.8. Map of identified animals (matched polygons) on site 3 using one standard deviation threshold (0.6) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m². No animal was identified on this site.

Site 4

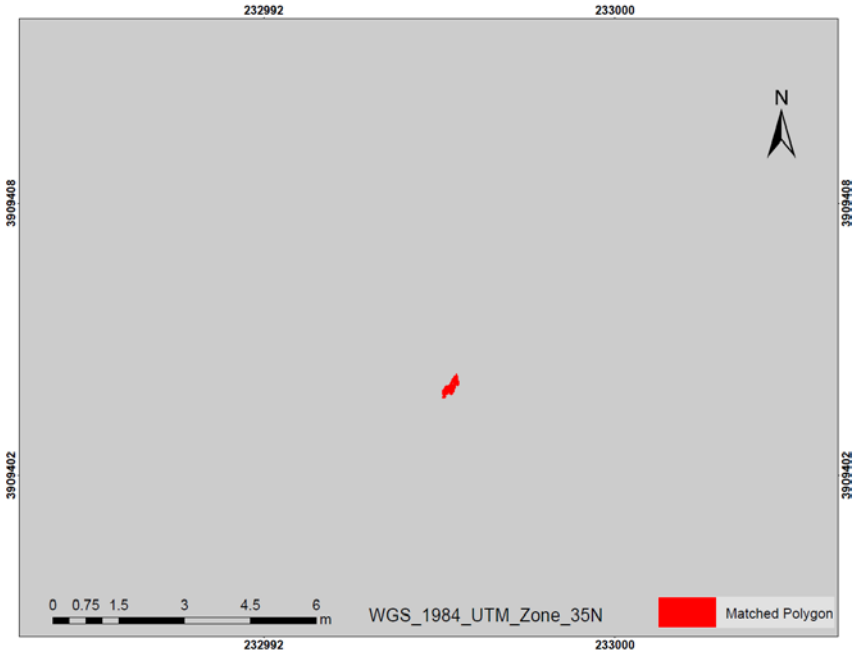


Figure 3.9. Map of identified animals (matched polygons) on site 4 using one standard deviation threshold (0.1) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m²

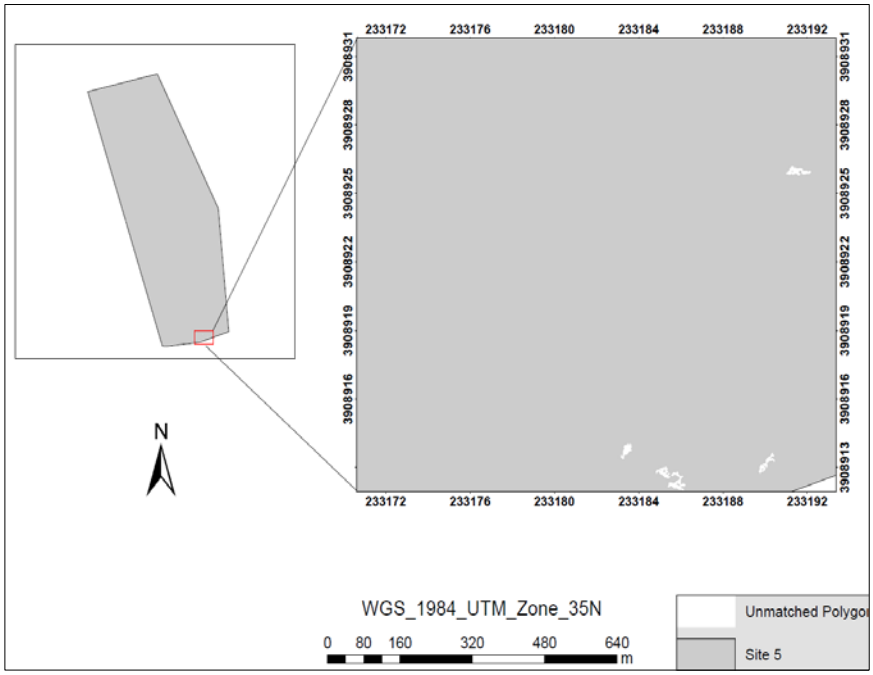


Figure 3.10. Map of identified (matched polygons) animals on site 5 using one standard deviation threshold (0.1) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m². This large site covered all the small test sites 1-4.

According to the figures above image differencing and vectorization generated eight polygons representing eight animals. Site 5 which was a mosaic of the whole site did not have polygons representing animals. It appears small sites had better results than large sites.

A table was generated representing true positives, false positives and false negatives. The results are summarized in Table 3.4. The result shows more detection was realized from manual count than the semi-automated method. There were also more FP than FN.

Table 3.3. Table representing TP, FP, and FN.

Site	Ground Count	Manual Count	True Positives (TP)	False Positives (FP)	False Negatives (FN)
1	3	3	3	94	0
2	8	7	4	14	4
3	3	3	0	4	3
4	4	3	1	0	2
5(large site)	18	16	0	9	15

3.3. Point cloud Comparison

Cloud compare open source software was used to compare two-point cloud data taken at intervals. Site 2 was selected for this comparison to test whether elevation data can be used in change detection. Site 2 was selected because it had the highest number of animals. First alignment accuracy of the pairs was assessed, and the RMSE result was 0.02. Using Cloud comparison of point cloud to point cloud the output generated is shown in Figure 3.11 as well as the mean (0.03) and standard deviation (0.11).

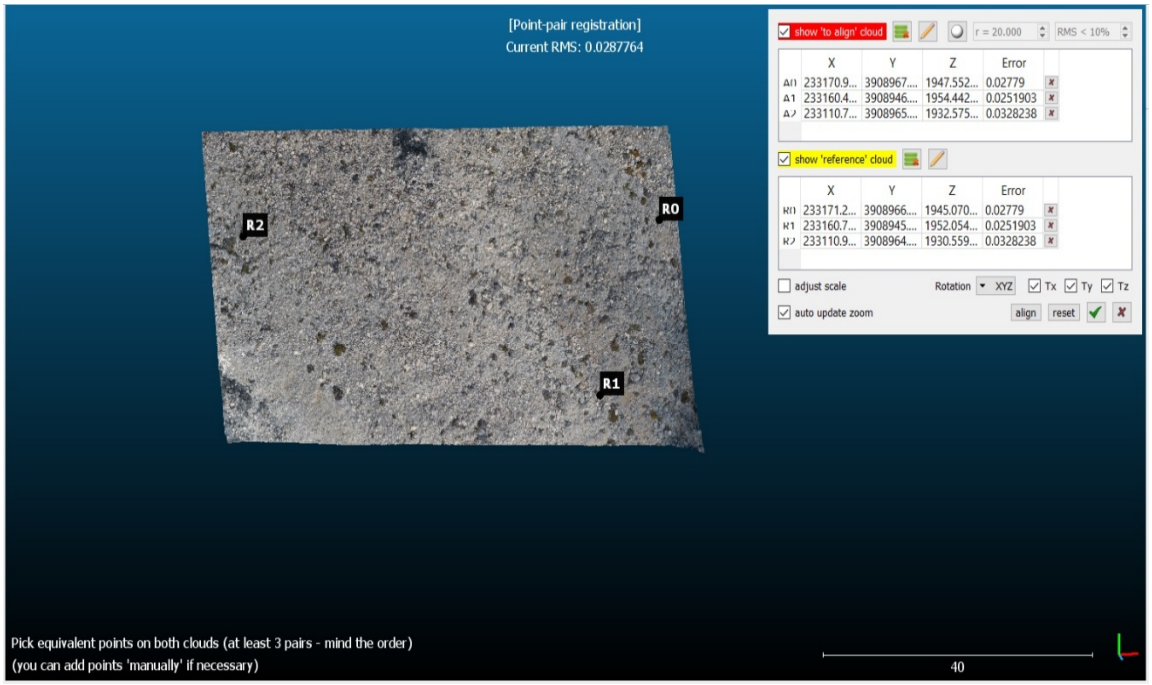


Figure 3.11. point cloud alignment and comparison

Cloud distance above the standard deviation represented change as shown in Figure 3.15. The figure shows change areas which were compared to image counts and 7 change areas were identified out of a possible 8 based on the ground counts. (Figure 3.12)

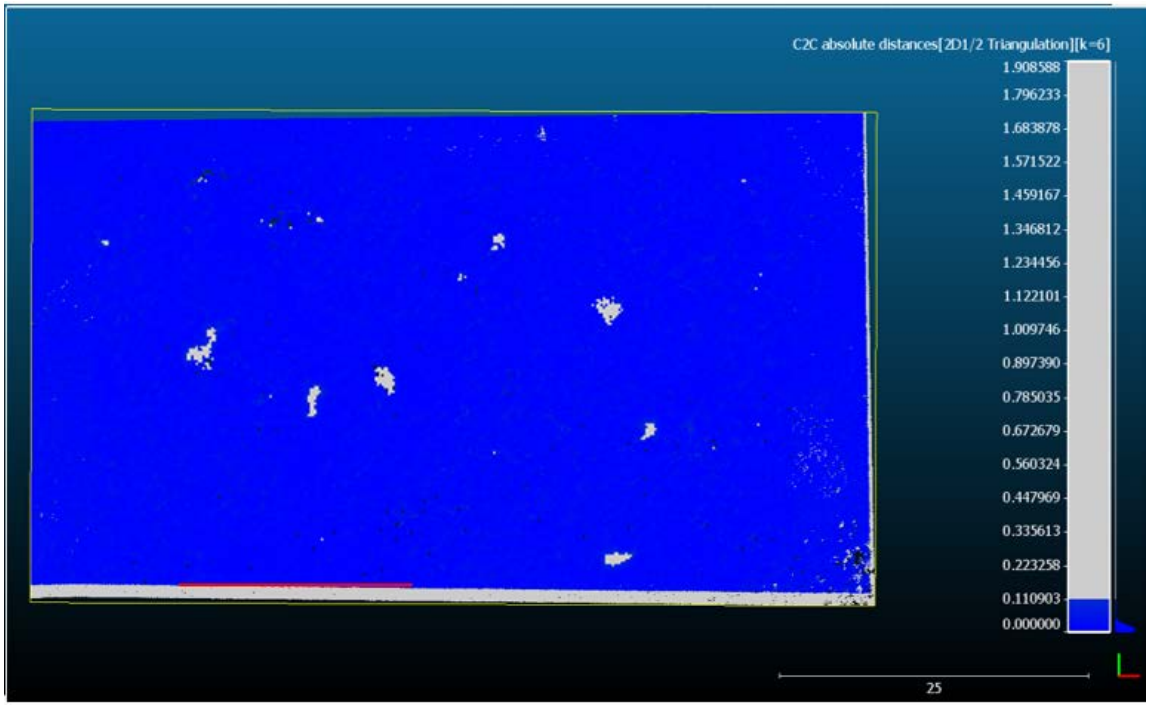


Figure 3.12. Results from point cloud comparison. 6 animal locations were identified. Some animals were merged in the process.

3.4. Image differencing using fusion approach (RGB+nDSM)

Based on the point cloud comparison results, nDSM was generated by subtracting DTM from DSM for each selected site. To test the feasibility of using a fusion of RGB and elevation layer, image differencing was done, and manual counts, as well as semi-automatic counts results, were recorded. Automatically generated polygons by vectorizing the differenced raster images were then compared with manual and ground counts to generate TP, FN, and FP as shown in Figure 3.16 to 3.20.

Table 3.4. Standard deviations for fusion approach selected sites

Site	Mean	Standard Deviation
1	3	13
2	-9	12
3	1	8
4	0.2	10
5	0	13

Table 3.5. Fusion approach results of TP, FP, FN.

Site	Ground Count	Manual Count	True Positives (TP)	False Positives (FP)	False Negatives (FN)
1	3	3	2	0	1
2	8	7	5	1	3
3	3	3	2	1	1
4	4	3	0	1	3
5(large site)	18	16	7	1980	8

Figure 3.16-3.20 shows results from image differencing process. Most of the animals were detected on the selected sites.

Fused image *Image difference results*

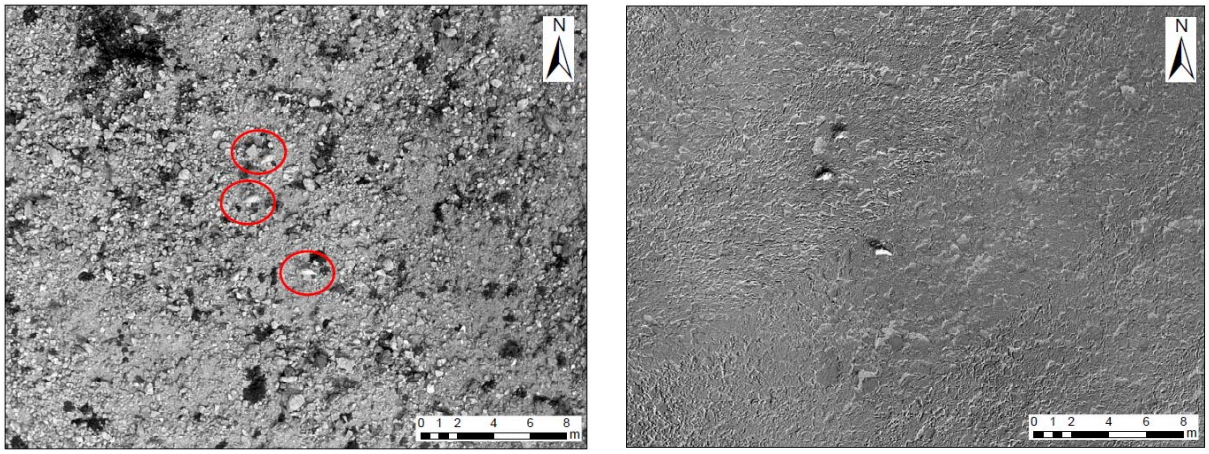


Figure 3.13. Site 2 fused original fused image (left) and difference result (right). Red shape indicates animal location

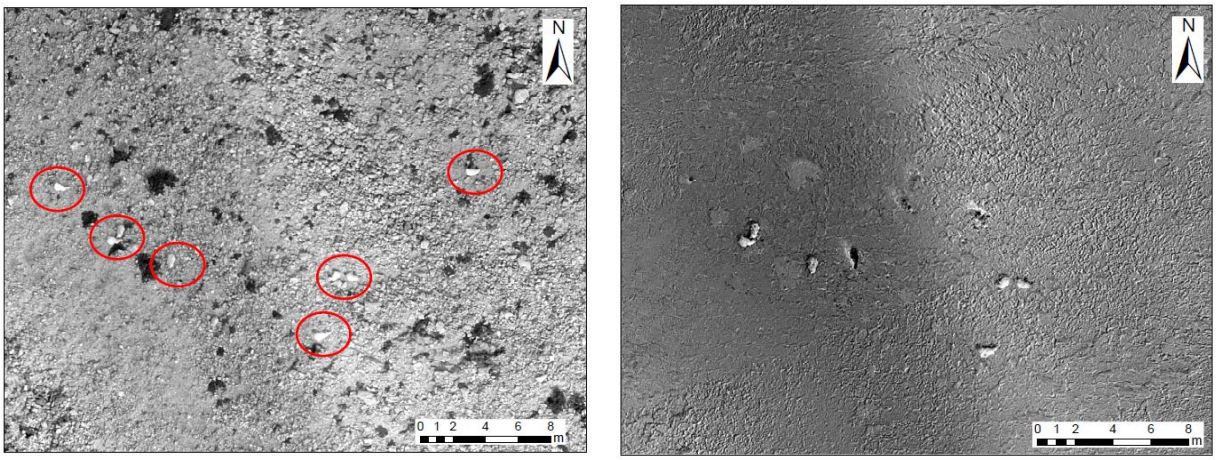


Figure 3.14. Site 2 fused original fused image (left) and difference result (right). 2 animals on the original image were not identified on the differenced image because they did not move thus no difference between the images on those locations.

A total of 9 polygon representing animals were realized in site 1-4 although site 4 had none. Site 5 representing the large site had 7. (Figure 3.25.). Interestingly, some animals that were not detected in site 4 were detected in the large site 5 which was a mosaic of the whole study site. Figure 3.22-3.25 presents maps of identified animals based on the polygons that were generated. Unmatched polygons represented falsely identified polygons as animals

Site 1

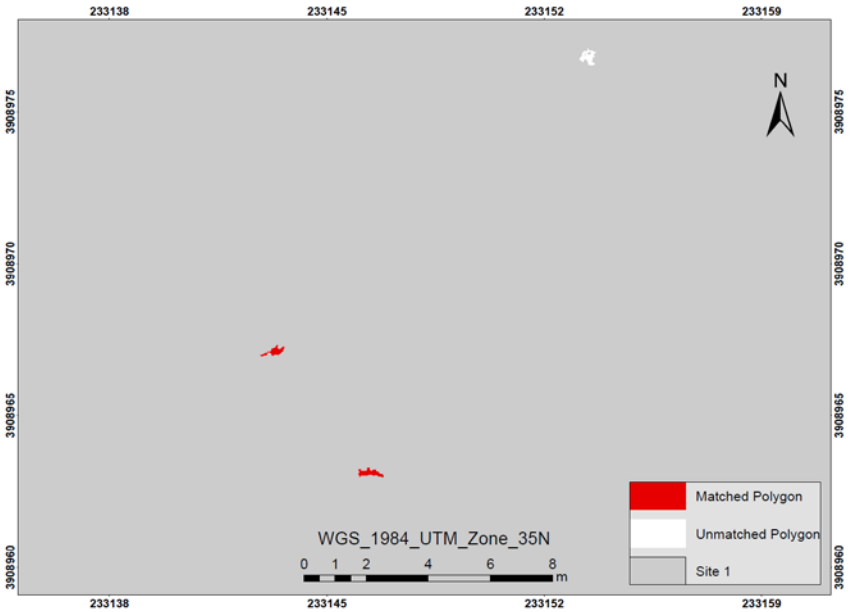


Figure 3.15. Map of identified animals (matched polygons) on site 1 using one standard deviation threshold (13) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m²

Site 2

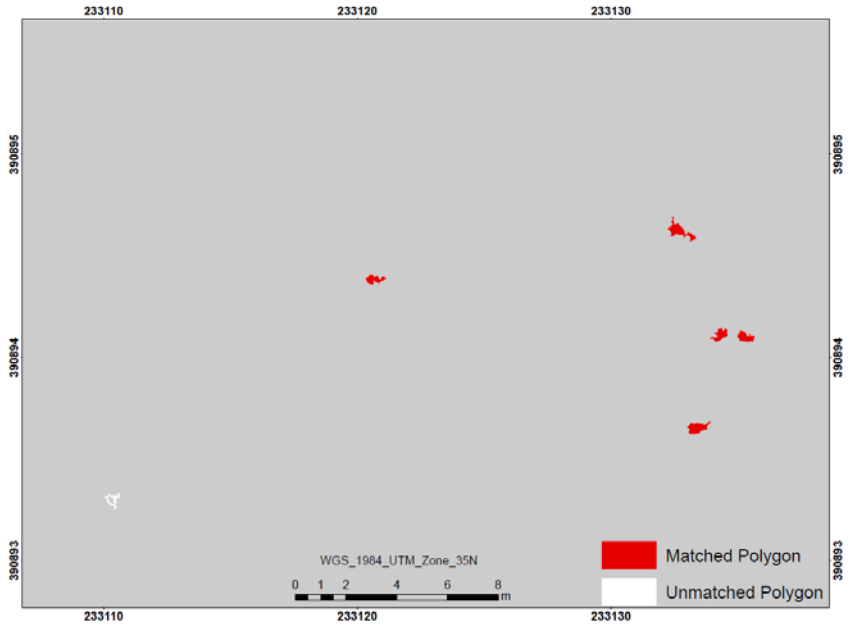


Figure 3.16. Map of identified animals (matched polygons) on site 2 using one standard deviation threshold (12) and animal size threshold of ≥ 0.1m² and ≤ 0.3m²

Site 3

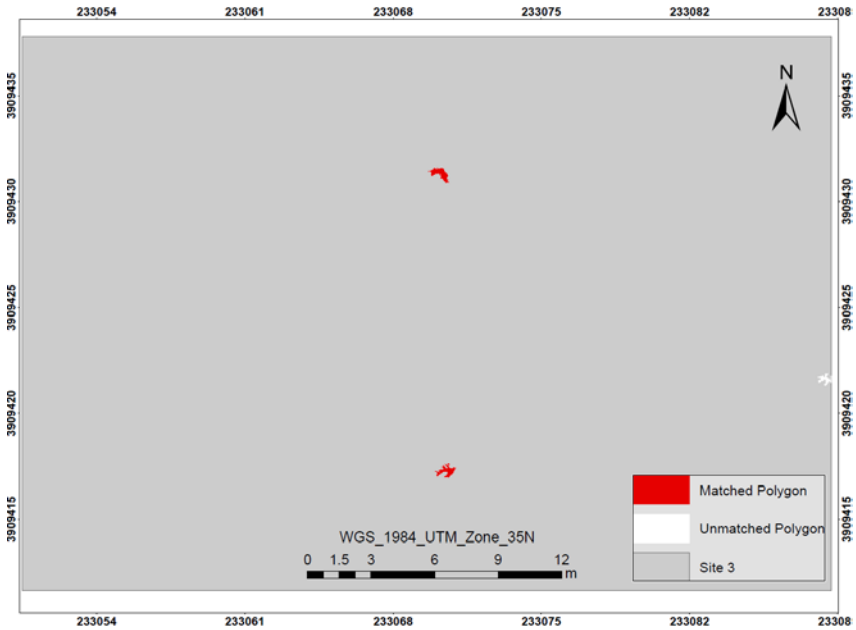


Figure 3.17. Map of identified animals (matched polygons) on site 3 using one standard deviation threshold (8) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m²

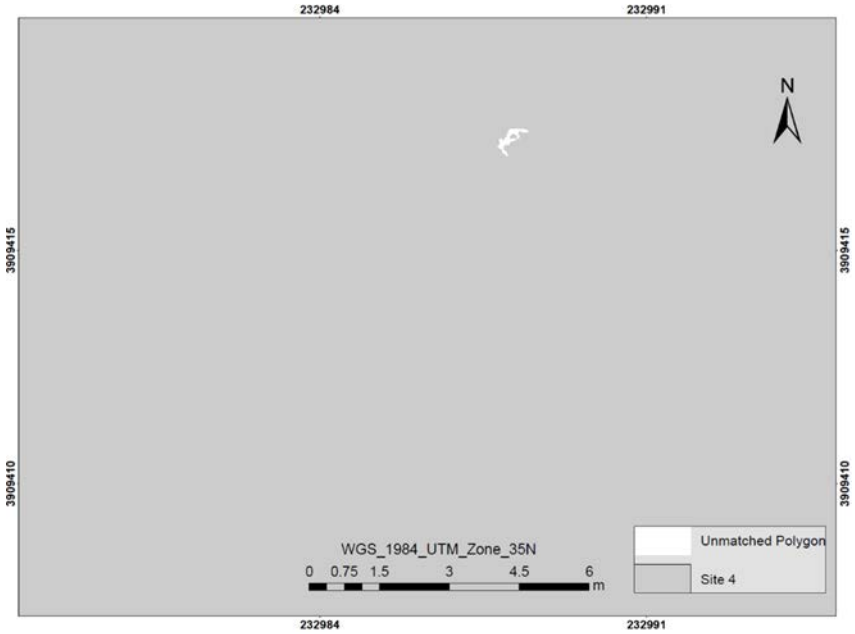


Figure 3.18. Map of identified animals (matched polygons) on site 4 using one standard deviation threshold (10) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m²

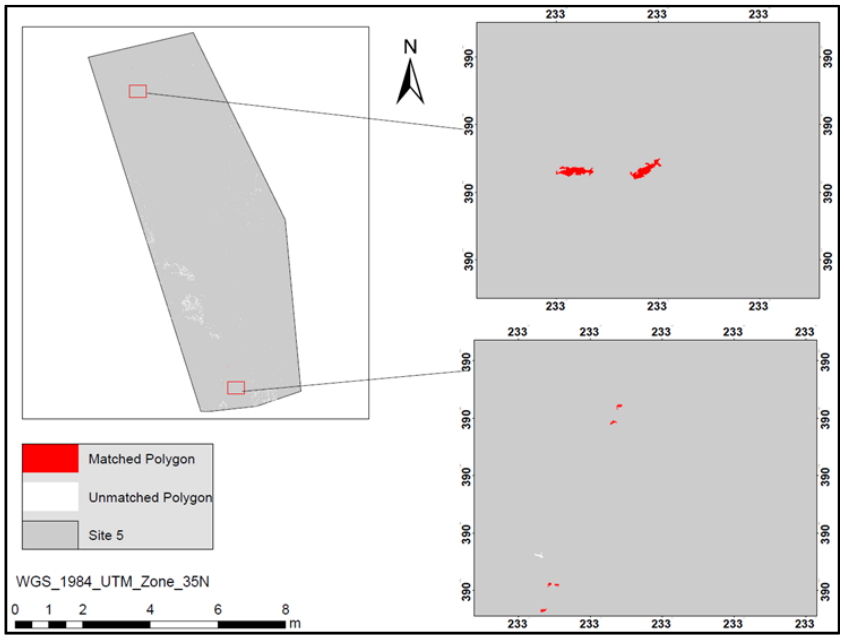


Figure 3.19. Map of identified animals (matched polygons) on site 5 using one standard deviation threshold (13) and animal size threshold of between ≥ 0.1m² and ≤ 0.3m².

3.5. Feasibility assessment of the image differencing method.

An assessment was done on both the manual count method and semi-automatic method. For the manual method, a percentage score was generated based on the results from ground counts which were compared with results from the original and differenced images. Precision and recall method was used to assess the accuracy and feasibility of the semi-automatic method and harmonized using an F score.

3.5.1. Manual Count

The overall detection rate using the original image and differenced image were 11% and 89% respectively. The computation was based on the results from the tables below. For the original image because of low contrast, accurate identification of animals was limited leading to a low overall detection rate.

Table 3.6. Manual detection results for the original and differenced image.

Site	Ground Count	Original Image	Differenced Images	% Detection rate (Original Image)	% Detection rate (Differenced image)
1	3	0	3	0	100
2	8	2	7	25	88
3	3	0	3	0	100
4	4	0	3	0	75
Total	18	2	16	11.1	89

3.5.2. Semi-Automatic Method

The formula for calculating precision, recall and F score is given below;

$$Precision = \frac{TP}{FP + TP} \qquad Recall = \frac{TP}{FN + TP} \qquad F\ score = 2\frac{P * R}{P + R}$$

Equation 3.1. Precision Recall and F score.

RGB image differencing approach

Site 2 had the highest precision and recall value at 100% and 50% respectively while the remaining sites had very low precision and recall values. The average precision and recall for the 4 sites were 6% and 41% respectively. The harmonic mean (F score) was 11%. It means, therefore, there was 41% correct detections achieved with a 6% precision.

Table 3.7. RGB approach

Site(0.1ha)	True Positives (TP)	False Positives (FP)	False Negatives (FN)	Precision	Recall	F Score
1	3	94	0	0.03	1.00	0.06
2	4	0	4	1.00	0.50	0.67
3	0	4	3	0.00	0.00	0.00
4	0	5	3	0.00	0.00	0.00
Mean	1.75	25.75	2.5	0.06	0.41	0.11

Fusion approach (RGB+ nDSM)

Fused images generated higher precision and recall as shown on the Table 3.9. Site 1-4 had 81%,71%, 67%, and 0% F measure respectively. Furthermore, the overall score (precision and recall) was 53% correct detections with a 75% precision. For site 4, both UAV RGB and fusion approach had 0% score implying no animals were visible.

Table 3.8. Fusion approach

Site(0.1ha)	True Positives (TP)	False Positives (FP)	False Negatives (FN)	Precision	Recall	F Score
001	2	0	1	1.00	0.67	0.80
002	5	1	3	0.83	0.63	0.71
003	2	1	1	0.67	0.67	0.67
004	0	1	3	0.00	0.00	0.00
Mean	2.25	0.75	2	0.75	0.53	**0.62**

Large area(9ha)

Results were also generated from the large test area measuring 9ha. Generally, the area had low precision and recall values as well as the F score. The fused image produced better results than RGB images for this area at 1% and 47% respectively (Table 3.10.) The F score was 2%.

Table 3.9. UAV RGB image and fusion approach for large site results

Site(9ha)	True Positives (TP)	False Positives (FP)	False Negatives (FN)	Precision	Recall	F1 Score
Site 5(RGB)	0	9	15	0	0.00	0.00
Site 5(Fused)	7	811	8	0.01	0.47	0.02

3.5.3. RGB image and fusion approach comparison.

The comparison was based on the overall precision, recall and F score rates for the calculated results. Generally, the fused images had a higher precision, recall and F score. Figure 3.20. shows Fusion and RGB image approach comparison for small sites.The F score for RGB image and fusion approach were 11% and 62% respectively. This implies RGB image approach had 41% correct detections achieved with an 11% precision while fusion approach had 53% detections achieved with a 75% precision.

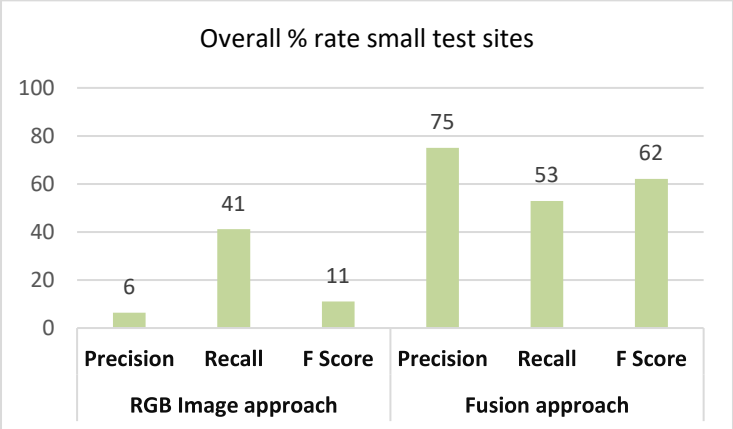


Figure 3.20. Fusion and RGB image approach comparison for selected small sites.

Comparison of the large site also showed a higher rate for the fusion approach although the overall scores were low (Figure 3.20). In summary, the fusion approach had 47% correct detections achieved with a 1% precision while RGB image approach did not have any results as presented in Figure 3.21.

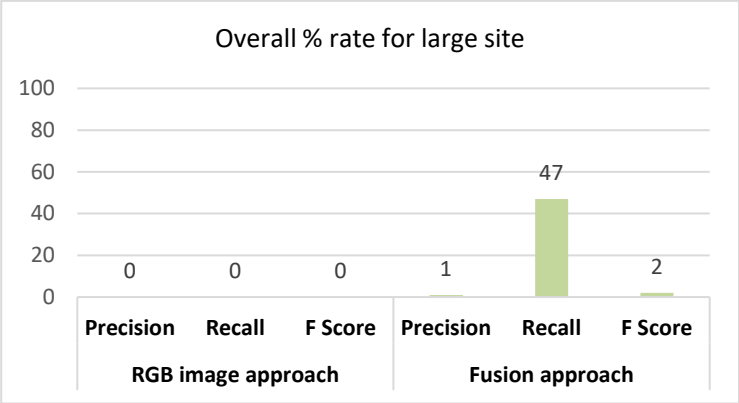


Figure 3.21. Fusion and RGB Image approach comparison for the large site.

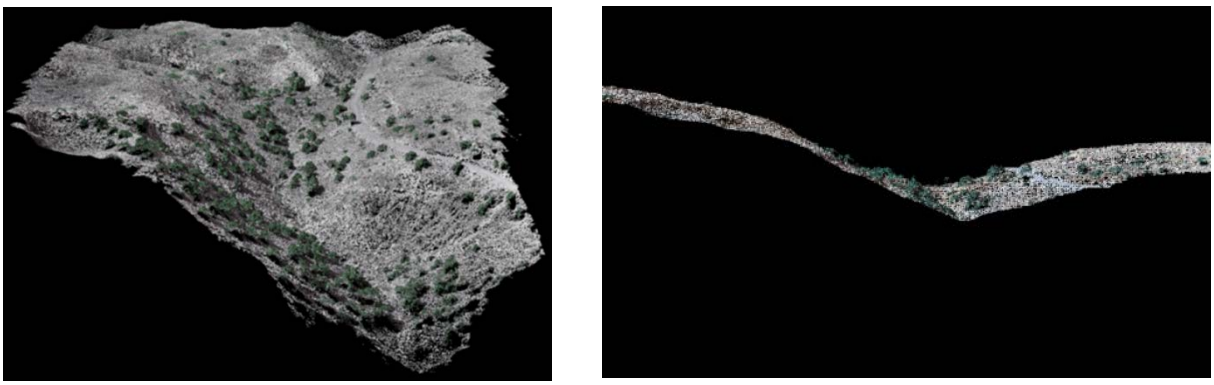
4. DISCUSSION

4.1. Using motion images

Flight records plays a crucial role in the determination of the 4 corner coordinates and subsequent georeferencing of a motion image. The overall RMSE was 15m. The assessment error is comparable to a result obtained by Barber et al. (2006) in their study on vision based target geolocation using a UAV which had an estimated error ranging between 20-40m. Even their focus was unrelated to animal detection, it still provided similar accuracy assessment. Eugster & Nebiker, (2008) also generated a real time digitizing accuracy of 6-15m in their study on georeferencing and integration of UAV video imagery.

The quality of the output depends on the used flight attitude, data and the level of synchronization between the video and flight data (Eugster & Nebiker 2008) Generally; the error was high and several reasons could have contributed to the high error. During a flight, a feature in motion may produce small sensor errors which are magnified during motion and geolocation. For instance, the motion images were taken at 45 degrees angle, and this encourages relief displacement effects. This method as earlier mentioned does not produce rectified images, therefore, the results from the motion image projected onto the orthorectified reference image do not precisely fall on the targeted corresponding feature point but presents an estimate of the location.

A generally moderate to flat topography has a better accuracy even without using DEM. Elevation data enhances the positional accuracy of features plus graphics on both the video and the map but, the area has a rough terrain which further explains the high RMSE (Figure 4.1.).



Site Image Side view

Figure 4.1: A dense point cloud image of the large site area showing the relief and the side view.

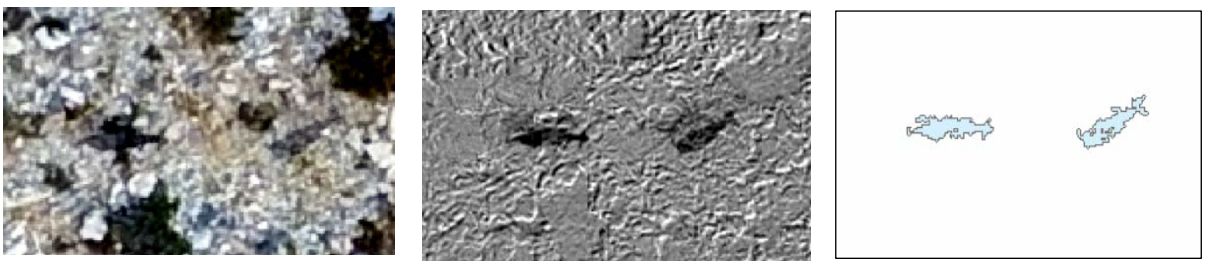
Despite the high error, the result is reliable since the animal species that were digitized were falling within a radius of 15m which is still good enough for mapping the location of goats and sheep. Eugster & Nebiker, (2008) reported that the geolocation accuracy of between 6-15m for video georeferencing method is adequate for many real-world applications although other applications might require precise accuracy. He further stated that the accuracy between 6-15m is applicable anywhere on the globe. A study done by Evangelou et al. (2014) in Northern Greece found that goats and sheep on average travel distances of 9km and 7.5km respectively per year although animals in a barren landscape travel for longer distances than the reported distances.

This method can therefore be helpful in animal abundance and distribution study, especially when conducting pre-mission analysis. The success of animal detection can be actualized if raw estimates of the

area of the target feature position in the motion image is known. Direct georeferencing method from motion video when viewed from cost-saving approach and simplicity, generates not only position information but also additional background information using the video frames of the geolocated areas. Furthermore, it is a source of reference, especially when verifying results from either manual or semi-automated approach in animal identification and enumeration. A big advantage of this method is that inaccessible areas can be reached, especially when random sampling and further modelling is of interest for instance Environmental Niche Modelling (ENM)

4.2. Short time interval Image differencing

The core of this study lies on the short-time image differencing on both an RGB image and RGB with elevation information acquired from a single sensor during two flight missions. The short time interval image differencing method generally improves the visibility and detection of the animals within a low contrast environment (Figure 4.2.). Moreover, it does not require extensive knowledge of the area by the user because the resultant output of animals is easily recognizable. The manual count accuracy rate was 89%. The fusion of ICA-transformed image and height information significantly improved the semi-automatic detection and counting in the selected sites compared to using RGB information only. This was based on the overall score at 12% using RGB images only and 64% using Fused RGB image + nDSM.



RGB image | differenced image | after cleaning process.

Figure 4.2. Image differencing and semi-automated process. The first column shows an image with low contrast, the second column shows a differenced image showing two visible goats and the third column shows the same goats after cleaning process.

Generally, this approach is comparable to an approach by (Terletzky & Ramsey 2014) who applied short time interval image differencing using Principle Component Analysis focusing on large animals(cattle and horses) and had a manual count of 82% with a mean and commission error of 53% and 18% respectively. It is also similar to Prošek & Šímová, (2019) study which applied fusion of RGB image and vertical information from a single UAV sensor to increase classification accuracy in shrubland vegetation. In his study, the overall accuracy was 88.2% for fusion approach and 73.3% for RGB image approach. A study by Husson et al. (2017) using similar fusion method improved the classification accuracy by 30%. Fused image is more informative than original image due to additional uncorrelated spectral information therefore doubles the data for detecting animals.

The robustness of the accuracy assessment is how well the application qualifies and quantifies the detection and enumeration results which in this study was measured using F score. Even though the recall has more weight in animal abundance estimation than precision, the latter is still useful in assessing the effectiveness of the method (Rey et al. 2017). For instance, in the study, 41% correct detections were realized with a precision of 6% for the optical image while 53% of correct detections where realized with a precision of 75% (Figure 4.3.)

The same results were also visible on the small test sites which generated high precision and recall rate. However, for the large area(0.9ha), fusion approach had 47% correct detections achieved with a 1%

precision while RGB image approach did not have any results. (Figure 4.3.). The small test sites demonstrated better results compared to the large site due primarily to less false positives.

Generally, accuracy is determined by the drone type, the sensor, flight attitude, texture of the surface, GCPs and the processing software. A possible explanation could be that as size increases, sensor errors both systematic and others such as stability, weather, wind and elevation and ground conditions are propagated. Flight movement disturbance due to the wind and small turbulence in some sections and input parameters such as flight speed can cause blur effect, and these blurs have a negative influence on the accuracy of the resultant images during pre-processing. The errors are propagated to the final results even when semi-automated methods have been used (Sieberth et al. 2013). Variability in scene illumination between adjacent images as well as two image pairs taken on the same day are a source of error which is multiplied when large image sets are used. Another major error includes image alignment due to insufficient orthorectification when matching two image pairs especially when ground control points are not used. The high spectral variability, image alignment coupled with blurred effect is, therefore, a major determinant of TP, FP, and FN.

Blurred effect affects the accuracy of terrain reconstruction from point cloud especially for rough terrain, and this has a trickle effect when using vertical information in the fusion approach. This is a probable explanation why the large test site for fused image had many false positives at 1,980 compared to RGB image approach which had 9 FP. According to Hollings et al., (2018) having a higher spectral and spatial resolution can affect image alignment which can generate small misregistration in overlaid images and this has the potential of increasing producer errors. When height information was added to optical image, there was an increase in FP especially in rugged and steep areas which have a high propensity for a mismatch.

Additionally, the photogrammetric point cloud processing relies on Structure from Motion (SfM) which is sometimes limited by gaps when key point features are not matched across image sequence. Studies have shown that LiDAR based point cloud is more accurate than photogrammetric point cloud because they are dense and insensitive to occlusion (Sankey et al. 2017). The result, therefore, limits the accuracy of image fusion and subsequently the differencing image process. However short time image differencing has the advantage of being applicable in both manual and semi-automated method to detect changes compared to other machine learning methods such as supervised and unsupervised classification. Despite challenges, it appears when height information is added to UAS RGB images it improves detection results, especially for relatively flat or gentle topography.

Manual counts although a tedious process when looking at large areas generated high correlation results with ground counts using image differencing output compared to the original RGB image. The accuracy was 89% to 11%. This is because terrain variability or micro- and macro- topographic variability made it difficult to differentiate animals from the background. Although manual counting on RS image is error-prone especially for large areas, using Image differencing improves the counts especially for areas of low animal density such as high desert rangelands.

The feasibility of the semi-automated method largely depends on the quality and quantity of the detections. Whereas recall has a higher weight in this study because it determines the quantity detected, an F measure was used to harmonize the recall and precision. The study did not apply weight on precision and recall, but the assumption was made that both are equal in weight. Even though F score introduced an element of biasness because it assumed both precision and recall have equal weight, the result, however, was a good measure on the viability and feasibility of the method. The low accuracy on the large test site

could be as a consequence of systematic errors in camera calibration due to non-usage of GCPs resulting into block deformation.

Interestingly, using manual method, both RGB image and fusion approach had a percentage accuracy of 89%. This implies that the addition of vertical information did not degrade or improve results of manual counts from differenced image output. Incorporation of both manual and semi-automatic method is a viable option for a survey of large areas. For instance, recall shows the number that was correctly detected therefore the undetected animals can be identified by using manual method thereby achieving a higher detection rate. Since the recall rate was 53 %, the 47% undetected can be compensated using the manual method. This reduces the time and labor-intensive nature of the manual method. Short time image differencing method however depends much on animal movement. If animals do not move, then they cannot be easily identified. 2 animals in site 2 were not identified because the animas were resting during UAV image acquisition.

The decision to use a combination of the methods depends on the topography, environment and animal density. This information can be achieved through a pre-mission survey using motion images. Estimates from the motion image georeferencing method form the basis for analysis of the final output. Although the concept is still new and requires further research and improvement in the algorithm for detection, the methodology is especially useful for inaccessible areas where ground truth data for verification is challenging to acquire. Even though the scope of the study was limited to detection and counts of animals, the results are a first step towards animal abundance estimation which requires other additional parameters. A combination of motion georeferencing, manual and semi-automatic methods as per this study can be feasible in an open rangeland environment with irregular terrain and low to medium animal density.

4.3. Limitations

One major limitation was the ground truth data acquisition. The flight mission relied on internal navigation parameters of the drone rather than Ground Control Points generated using differential GNSS device and ground station. This limited the accuracy of the photogrammetric process for all the products that were used. Certain areas had a blurred effect. Currently new RTK enabled UAS have been introduced in the market.

The choice of date for data collection is another important factor, for this study field work was done during the month of September 2018 when temperatures had already started to decrease. As a consequence of the early onset of the cold season most of the animals had already migrated from the high mountain and moved towards the lowlands. More animals within the study area would be better to form a strong basis for the detection methods studied.

Only one large site was chosen for detailed study out of the 6 sites that had been pre-selected. The site was chosen because of the presence of animals. The outcomes of this study would have been better if all the 6 sites were studied to factor in altitude, terrain difference, and type of vegetation. Limited visibility due to rugged terrain and battery life prevented surveying of large areas. Furthermore, regulations prevent flying farther than 500m away.

5. CONCLUSION AND RECOMMENDATION

5.1. Conclusion

UAV imagery offers new frontier in ecological research and studies that have been done so far albeit on a proof of concepts are demonstrating the potential benefits of this technology in animal detection and counting. Previous researches done have not focused on open rangeland areas with low contrast and extreme topography which have myriads of challenges in investigating. This study investigated the possibilities of detection of animals using motion image, point cloud, and orthophoto image by applying manual and semi-automatic techniques. The main aim was to assess these UAS data products and their usefulness in animal detection. This is a first step towards animal abundance estimation and understanding of the drivers of animal movements. This in turn could be linked with stocking rates and spatial representation of grazing pressure patterns. This would only form a part in the puzzle of rangeland management that requires other parameters such as carrying capacity, biomass estimation, biomass composition, and soil conditions (which were outside the scope of this study).

The findings of this study indicate that a combination of fusion of RGB and height information can significantly improve the accuracy of animal detection and counting. It also suggests that short time image differencing is effective for manual counting and requires little user knowledge of the area. For semi-automated counting, it is feasible for a small area. The motion videos geolocation method is useful for pre-mission estimation. The findings further indicated that a combination of manual and semi-automated techniques improves animal survey useful for ecosystem indicators generation for management of rangelands.

Using the semi-automated approach in estimation still, need further research especially for large areas. If precise and accurate animal detection is achieved from UAV images in a semi-automated way, this will lead to improved abundance estimates. Development of D-RTK-GNSS, a high precision navigation and positioning system, and upgrading of UAV images processing software such as Pix4D to use such data is a major step towards achieving better surveys. Despite using low cost drone and lack of GCPs, the results are promising in rangeland management and a step towards the integration of both manual and semi-automatic method.
Below is a synopsis of the answers to the research questions of this study.

Can motion images provide high accuracy in spatial location and distribution of animals?
The study revealed that the RMSE was 15m which is within the global accuracy results between 6-15m based on other studies that have been done. Therefore, even though the accuracy is not high enough, it is reliable for species location and distribution study especially for a steep and rugged area and when doing a pre-mission. This method is, therefore, suitable if we are more interested in numbers and not exact location.

Can image differencing using Independent Component Analysis provide overall detection rate comparable to a ground-based survey in the detection of animals like goats and sheep in open rangeland?
For manual counting method, image differencing provided a high correlation with the ground-based survey (89%) therefore it is feasible, but for the semi-automated method, the result is dependent on the size of the area under investigation. High F score is visible in small sites for fused images for example site 2 had an F score percentage rate of 32% for optical and 77% for fused images. The large site(fused) had

an accuracy score of 2%. Generally, for the large sites, the semi-automated method does not provide a reliable detection rate comparable to the manual and ground-based survey.

Can image fusion improve the detection of animals in open rangeland?
The study indicate that integration of RGB and height information could improve the detection rate accuracy of the semi-automated method. For manual count using a differenced image, image fusion did not improve the overall result.

Are the accuracy rates high enough to consider the feasibility of the image differencing method?
Generally, the rates for the manual method on the differenced image are high enough for consideration in ecological research in open rangeland. Although hidden animals cannot be easily detected, the overall percentage estimation is high enough for ecosystem indicators generation but there are other sources of errors, for instance when the animal move, they might be captured more than one or not at all under a certain flight plan. For the semi-automated method, the rates are not high enough for wide usage in open rangeland. The false positives increased with increase in the size of the area under study and this influenced the overall result. It is clear the detectability is high enough for the manual count but not high enough for semi-automated count, especially on large areas.

5.2. Recommendation

With the continued rapid advancement in the use of UAV image in remote sensing, advanced software, a complex algorithm and machine learning both for object-based and pixel-based analysis, animal abundance and inventory estimation will increase considerably. Applying a combination of manual and semi-automated approach in given rangeland is feasible. If the recall rate is high enough in a given setting, manual counting method using image differenced image can be used to compensate.

The research focused more on open rangeland and overcoming terrain and low contrast characteristics and, this is a step towards studies on animal detection in vegetated areas with trees. For the free-ranging animals, if complete survey is not possible, inputs from similar work can help estimate stocking rates. Furthermore, animal detection and counts can serve as input to Environmental Niche Models (Species Distribution Models). ENMs can further pinpoint the factors controlling animal movement, and map grazing pressure patterns. This is crucial in rangeland management as well as conservation in protected areas such as the Lefka Ori.

For future research, LiDAR data can be fused with RGB images, and the results analyzed on its applicability as well as its comparability to Photogrammetric point cloud. Future research should also look at the possibility of using thermal sensors in UAV or coupling of thermal sensors and RGB for detection. Object-based method approach in detection should also be explored. While most research has focused on homogenous and small areas with high contrast, further research is strongly recommended in larger heterogenous areas with low contrast. Data collection should be done when animals are not migrating preferably in June and July. Also, its recommended that the time of acquisition should be mid-day to limit the effects of shadows and sun illumination angle. D-RT-GNSS enabled UAS should also be used.

LIST OF REFERENCE

Barber, D. B., Redding, J. D., McLain, T. W., Beard, R. W., & Taylor, C. N. (2006). 'Vision-based target geo-location using a fixed-wing miniature air vehicle', *Journal of Intelligent and Robotic Systems: Theory and Applications*. DOI: 10.1007/s10846-006-9088-7

Benassi, F., Dall'Asta, E., Diotri, F., Forlani, G., Cella, U. M. di, Roncella, R., & Santise, M. (2017). 'Testing accuracy and repeatability of UAV blocks oriented with GNSS-supported aerial triangulation', *Remote Sensing*, 9/2: 1–23. DOI: 10.3390/rs9020172

Briassoulis, H. (2003). 'Crete: Endowed by nature, privileged by geography, threatened by tourism?', *Journal of Sustainable Tourism*, 11/2–3: 97–115. DOI: 10.1080/09669580308667198

Duro, D. C., Franklin, S. E., & Dubé, M. G. (2012). 'A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery', *Remote Sensing of Environment*, 118: 259–72. DOI: 10.1016/j.rse.2011.11.020

Eugster, H., & Nebiker, S. (2008). 'UAV-Based Augmented Monitoring–Real-Time Georeferencing and Integration of Video Imagery with Virtual Globes', *Drone*, 37: 1229–36.

Evangelou, C., Yiakoulaki, M., & Papanastasis, V. (2014). 'Spatio-temporal analysis of sheep and goats grazing in different forage resources of Northern Greece', 13/1: 205–13. DOI: 10.2478/hacq-2014-0001

Fernández-Calzado, R., Ghosn, D., Gottfried, M., Kazakis, G., Molero Mesa, J., Pauli, H., & Merzouki, A. (2013). 'Patterns of endemism along an elevation gradient in Sierra Nevada (Spain) and Lefka Ori (Crete, Greece)', *Pirineos*, 168/0: 7–24. DOI: 10.3989/Pirineos.2013.168001

Van Gemert, J. C., Verschoor, C. R., Mettes, P., Epema, K., Koh, L. P., & Wich, S. (2015). 'Nature Conservation Drones for Automatic Localization and Counting of Animals'. Lagapito L., Bronstein M., & Rother C. (eds) *Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science*, pp. 255–70. Springer, Cham. DOI: 10.1007/978-3-319-16178-5_17

Girardeau-Montaut, D., Roux, M., Marc, R., & Thibault, G. (2005). 'Change Detection on Points Cloud Data Acquired With a Ground', *International Society for Photogrammetry and Remote Sensing Workshop - Laser scanning*, 13/May 2014.

Gonzalez, L. F., Montes, G. A., Puig, E., Johnson, S., Mengersen, K., & Gaston, K. J. (2016). 'Unmanned aerial vehicles (UAVs) and artificial intelligence revolutionizing wildlife monitoring and conservation', *Sensors (Switzerland)*, 16/1. DOI: 10.3390/s16010097

Gormley, A. M., Forsyth, D. M., Griffioen, P., Lindeman, M., Ramsey, D. S., Scroggie, M. P., & Woodford, L. (2011). 'Using presence-only and presence-absence data to estimate the current and potential distributions of established invasive species.', *The Journal of applied ecology*, 48/1: 25–34. Wiley-Blackwell. DOI: 10.1111/j.1365-2664.2010.01911.x

Groom, G., Stjernholm, M., Nielsen, R. D., Fleetwood, A., & Petersen, I. K. (2013). 'Remote sensing image data and automated analysis to describe marine bird distributions and abundances', *Ecological Informatics*, 14: 2–8. Elsevier. DOI: 10.1016/J.ECOINF.2012.12.001

Grove, A. T., & Rackham, O. (1993). 'Threatened landscapes in the Mediterranean: examples from Crete', *Landscape and Urban Planning*, 24/1–4: 279–92. Elsevier. DOI: 10.1016/0169-2046(93)90107-O

Hao, L., Pan, C., Fang, D., Zhang, X., Zhou, D., Liu, P., Liu, Y., et al. (2018). 'Quantifying the effects of overgrazing on mountainous watershed vegetation dynamics under a changing climate', *Science of The Total Environment*, 639: 1408–20. Elsevier. DOI: 10.1016/J.SCITOTENV.2018.05.224

Hodgson, J. C., Mott, R., Baylis, S. M., Pham, T. T., Wotherspoon, S., Kilpatrick, A. D., Raja Segaran, R., et al. (2018). 'Drones count wildlife more accurately and precisely than humans', (N. Yoccoz, Ed.)*Methods in Ecology and Evolution*, 7/5: 1160–7. Wiley/Blackwell (10.1111). DOI: 10.1111/2041-210X.12974

Hollings, T., Burgman, M., van Andel, M., Gilbert, M., Robinson, T., & Robinson, A. (2018). 'How do you find the green sheep? A critical review of the use of remotely sensed imagery to detect and count animals', (J. McPherson, Ed.)*Methods in Ecology and Evolution*, 9/4: 881–92. Wiley/Blackwell (10.1111). DOI: 10.1111/2041-210X.12973

Husson, E., Reese, H., Ecke, F., Husson, E., Reese, H., & Ecke, F. (2017). 'Combining Spectral Data and a DSM from UAS-Images for Improved Classification of Non-Submerged Aquatic Vegetation', *Remote Sensing*, 9/3: 247. Multidisciplinary Digital Publishing Institute. DOI: 10.3390/rs9030247

Hyvärinen, A., & Oja, E. (2000). *Independent Component Analysis: Algorithms and Applications*. *Neural Networks*, Vol. 13. Retrieved January 14, 2019, from <https://www.cs.helsinki.fi/u/ahyvarin/papers/NN00new.pdf>

Jiménez-Hernández, H. (2010). 'Background subtraction approach based on independent component analysis.', *Sensors (Basel, Switzerland)*, 10/6: 6092–114. Multidisciplinary Digital Publishing Institute (MDPI). DOI: 10.3390/s100606092

Jin, L., Ying-cheng, L., De-long, L., Chang-sheng, T., & Wen-hao, Z. (2014). 'The orthorectified technology for UAV aerial Remote Sensing image based on the Programmable GPU', *IOP Conference Series: Earth and Environmental Science*, 17/1: 012-201. IOP Publishing. DOI: 10.1088/1755-1315/17/1/012201

Khan, M. B., Nisar, H., & Ng, C. A. (2018). 'Image Processing and Analysis of Phase-Contrast Microscopic Images of Activated Sludge to Monitor the Wastewater Treatment Plants', *IEEE Access*, 6: 1778–91. DOI: 10.1109/ACCESS.2017.2780249

Kumar, V., & Gupta, P. (2012). 'International Journal of Emerging Technology and Advanced Engineering Importance of Statistical Measures in Digital Image Processing', *International Journal of Emerging Technology and Advanced Engineering*, 2/8.

Laity, J. (2009). *Deserts and Desert Environments*. New York: John Wiley & Sons. DOI: DOI: 10.1111/j.1475-4959.2009.00347_6.x

Laliberte, A. S., & Ripple, W. J. (2003). 'Automated wildlife counts from remotely sensed imagery', *Wildlife Society Bulletin*. DOI: 10.2307/3784314

LaRue, M., Lynch, H. J., Lyver, P. O. B., Barton, K., Ainley, D. G., Pollard, A., Fraser, W. R., et al. (2014). 'A method for estimating colony sizes of Adélie penguins using remote sensing imagery', *Polar Biology*, 37/4: 507–17. Springer Berlin Heidelberg. DOI: 10.1007/s00300-014-1451-8

LaRue, M., Stapleton, S., Porter, C., Atkinson, S., Atwood, T., Dyck, M., & Lecomte, N. (2015). 'Testing methods for using high-resolution satellite imagery to monitor polar bear abundance and distribution', *Wildlife Society Bulletin*, 39/4: 772–9. Wiley-Blackwell. DOI: 10.1002/wsb.596

Lee, T.-W., & Lewicki, M. S. (2002). *Unsupervised Image Classification, Segmentation, and Enhancement Using ICA Mixture Models*. Retrieved January 14, 2019, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.324.2526&rep=rep1&type=pdf>

Lorent, H., Sonnenschein, R., Tsiourlis, G. M., Hostert, P., & Lambin, E. (2009). 'Livestock subsidies and rangeland degradation in central Crete', *Ecology and Society*, 14/2. DOI: 10.5751/ES-03229-140241

McEvoy, J. F., Hall, G. P., & McDonald, P. G. (2016). 'Evaluation of unmanned aerial vehicle shape, flight path and camera type for waterfowl surveys: disturbance effects and species recognition', *PeerJ*, 4: e1831. DOI: 10.7717/peerj.1831

McNeill, S., Barton, K., Lyver, P., & Pairman, D. (2011). 'Semi-automated penguin counting from digital aerial photographs'. *2011 IEEE International Geoscience and Remote Sensing Symposium*, pp. 4312–5. IEEE. DOI: 10.1109/IGARSS.2011.6050185

Menke, J., & Eric Bradford, G. (1992). 'Rangelands', *Agriculture, Ecosystems & Environment*. DOI: 10.1016/0167-8809(92)90024-6

Mirvakhabova, L., Pukalchik, M., Matveev, S., Tregubova, P., & Oseledets, I. (2018). 'Field heterogeneity detection based on the modified FastICA RGB-image processing', *Journal of Physics: Conference Series*, 1117/1: 012009. IOP Publishing. DOI: 10.1088/1742-6596/1117/1/012009

MISB. (2017). *Motion Imagery Standards Profile Motion Imagery Standards Board MISP-2018.1: Motion Imagery Handbook*. Retrieved December 18, 2018, from <http://www.gwg.nga.mil/misb/docs/misp/MISP-2018.1_Motion_Imagery_Handbook.pdf>

Powers, D. M. W., & Ailab. (2011). 'Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness & amp; Correlation', 2/1: 37–63.

Previtali, M., Barazzetti, L., Scaioni, M., Wallace, L., Lucieer, A., Watson, C., Turner, D., et al. (2012). 'UAV Photogrammetry', *Remote Sensing*. DOI: doi:10.3929/ethz-a-005939264

Prošek, J., & Šímová, P. (2019). 'UAV for mapping shrubland vegetation: Does fusion of spectral and vertical information derived from a single sensor increase the classification accuracy?', *International Journal of Applied Earth Observation and Geoinformation*, 75: 151–62. Elsevier. DOI: 10.1016/J.JAG.2018.10.009

Rango, A., & Laliberte, A. (2010). 'Rangeland resource assessment, monitoring, and management using Unmanned Aerial Vehicle-based Remote Sensing'. *2010 IEEE International Geoscience and Remote Sensing Symposium*, pp. 608–11. Honolulu: IEEE. DOI: 10.1109/IGARSS.2010.5651659

Rango, A., Laliberte, A., Herrick, J. E., Winters, C., Havstad, K., Steele, C., & Browning, D. (2009). 'Unmanned Aerial Vehicle-based remote sensing for rangeland assessment, monitoring, and management', *Journal of Applied Remote Sensing*, 3/1. DOI: 10.1117/1.3216822

Rey, N., Volpi, M., Joost, S., & Tuia, D. (2017). 'Detecting animals in African Savanna with UAVs and the crowds', *Remote Sensing of Environment*, 200: 341–51. Elsevier. DOI: 10.1016/J.RSE.2017.08.026

Sankey, T., Donager, J., McVay, J., & Sankey, J. B. (2017). 'UAV lidar and hyperspectral fusion for forest monitoring in the southwestern USA', *Remote Sensing of Environment*. DOI: 10.1016/j.rse.2017.04.007

Sarp, G., Erener, A., Duzgun, S., & Sahin, K. (2014). 'An approach for detection of buildings and changes in buildings using orthophotos and point clouds: A case study of Van Erriş earthquake', *European Journal of Remote Sensing*, 47/1: 627–42. Taylor & Francis. DOI: 10.5721/EuJRS20144735

Seymour, A. C., Dale, J., Hammill, M., Halpin, P. N., & Johnston, D. W. (2017). 'Automated detection and enumeration of marine wildlife using unmanned aircraft systems (UAS) and thermal imagery', *Scientific Reports*, 7: 45–127. Nature Publishing Group. DOI: 10.1038/srep45127

Shi, J., Wang, J., & Xu, Y. (2011). 'Object-Based Change Detection Using Georeferenced UAV Images', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38/1-C22: 177–82.

Shroder, J. F., Sivanpillai, R., Angerer, J. P., Fox, W. E., & Wolfe, J. E. (2016). 'Land Degradation in Rangeland Ecosystems', *Biological and Environmental Hazards, Risks, and Disasters*, 277–311. Academic Press. DOI: 10.1016/B978-0-12-394847-2.00017-6

Sieberth, T., Wackrow, R., & Chandler, J. H. (2013). 'Automatic isolation of blurred images from UAV image sequences', *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1/W2: 4–6.

Sykora-Bodie, S. T., Bezy, V., Johnston, D. W., Newton, E., & Lohmann, K. J. (2017). 'Quantifying nearshore sea turtle densities: Applications of Unmanned Aerial Systems for population assessments', *Scientific Reports*, 7/1: 17690. DOI: 10.1038/s41598-017-17719-x

Terletzky, P. A., & Ramsey, R. D. (2016). 'Comparison of Three Techniques to Identify and Count Individual Animals in Aerial Imagery', *Journal of Signal and Information Processing*, 07/03: 123–35. Scientific Research Publishing. DOI: 10.4236/jsip.2016.73013

Terletzky, P., & Ramsey, R. D. (2014). 'A Semi-Automated Single Day Image Differencing Technique to Identify Animals in Aerial Imagery', (C. Ricotta, Ed.)*PLoS ONE*, 9/1: e85239. Public Library of Science. DOI: 10.1371/journal.pone.0085239

Trathan, P. N. (2004). 'Image analysis of color aerial photography to estimate penguin population size', *Wildlife Society Bulletin*. DOI: 10.2193/0091-7648(2004)32[332:IAOCAP]2.0.CO;2

Vogiatzakis, I., & Rackham, O. (2008). 'Mediterranean island landscapes'. Vogiatzakis I., Pungetti G., &

Mannion A. M. (eds) *Landscape Series*, Landscape Series, Vol. 9, pp. 245–70. Springer Netherlands: Dordrecht. DOI: 10.1007/978-1-4020-5064-0

Yang, Z., Wang, T., Skidmore, A. K., de Leeuw, J., Said, M. Y., & Freer, J. (2014). 'Spotting East African mammals in open savannah from space', (M. Cristani, Ed.)*PLoS ONE*, 9/12. Public Library of Science. DOI: 10.1371/journal.pone.0115989

Yuan, J., Wang, D., Li, R., & Member, S. (2014). 'Remote Sensing Image Segmentation by Combining Spectral and Texture Features', *IEEE Transactions on Geoscience and Remote Sensing*, 52/1. DOI: 10.1109/TGRS.2012.2234755

APPENDICES

Appendix A. part of the Verbose CSV file generated from flight records.

	CUSTOM.updateTime	UNIX Time Stamp	SensorLatitud	SensorLongitud	SensorAltitude	PlatformHeadin	PlatformPitch	PlatformRoll	HorizontalFO	VerticalFOV	SensorRelati	SensorRelati	S
2	9/21/2018 12:56	1537531014550000	35.265217	24.087937	1581.000	161.7	-45	0	84	54	0	-45	
3	9/21/2018 12:56	1537531014754000	35.265213	24.087939	1581.000	161.7	-45	0	84	54	0	-45	
4	9/21/2018 12:56	1537531014858000	35.265205	24.087943	1581.000	161.7	-45	0	84	54	0	-45	
5	9/21/2018 12:56	1537531014961000	35.265201	24.087945	1581.000	161.7	-45	0	84	54	0	-45	
6	9/21/2018 12:56	1537531014961000	35.265197	24.087946	1581.000	161.7	-45	0	84	54	0	-45	
7	9/21/2018 12:56	1537531015064000	35.265194	24.087948	1581.000	161.7	-45	0	84	54	0	-45	
8	9/21/2018 12:56	1537531015167000	35.26519	24.08795	1581.000	161.7	-45	0	84	54	0	-45	
9	9/21/2018 12:56	1537531015268000	35.265186	24.087952	1581.000	161.7	-45	0	84	54	0	-45	
10	9/21/2018 12:56	1537531015371000	35.265181	24.087953	1581.000	161.6	-45	0	84	54	0	-45	
11	9/21/2018 12:56	1537531015473000	35.265177	24.087955	1581.000	161.6	-45	0	84	54	0	-45	
12	9/21/2018 12:56	1537531015574000	35.265173	24.087957	1573.816	161.6	-45	0	84	54	0	-45	
13	9/21/2018 12:56	1537531015677000	35.265169	24.087959	1573.816	161.6	-45	0	84	54	0	-45	
14	9/21/2018 12:56	1537531015779000	35.265164	24.08796	1573.816	161.6	-45	0	84	54	0	-45	
15	9/21/2018 12:56	1537531015880000	35.26516	24.087962	1573.816	161.6	-45	0	84	54	0	-45	
16	9/21/2018 12:56	1537531015983000	35.265156	24.087964	1573.816	161.6	-45	0	84	54	0	-45	
17	9/21/2018 12:56	1537531016290000	35.265152	24.087965	1573.816	161.6	-45	0	84	54	0	-45	
18	9/21/2018 12:56	1537531016395000	35.26514	24.08797	1573.816	161.6	-45	0	84	54	0	-45	
19	9/21/2018 12:56	1537531016497000	35.265136	24.087972	1573.816	161.6	-45	0	84	54	0	-45	
20	9/21/2018 12:56	1537531016600000	35.265132	24.087973	1573.970	161.6	-45	0	84	54	0	-45	
21	9/21/2018 12:56	1537531016706000	35.265128	24.087975	1566.706	161.6	-45	0	84	54	0	-45	
22	9/21/2018 12:56	1537531016808000	35.265124	24.087976	1566.706	161.6	-45	0	84	54	0	-45	
23	9/21/2018 12:56	1537531017014000	35.26512	24.087978	1566.706	161.6	-45	0	84	54	0	-45	
24	9/21/2018 12:56	1537531017117000	35.265112	24.087981	1566.706	161.6	-45	0	84	54	0	-45	
25	9/21/2018 12:56	1537531017219000	35.265108	24.087983	1566.706	161.6	-45	0	84	54	0	-45	
26	9/21/2018 12:56	1537531017322000	35.265104	24.087985	1566.706	161.6	-45	0	84	54	0	-45	
27	9/21/2018 12:56	1537531017424000	35.2651	24.087987	1566.706	161.6	-45	0	84	54	0	-45	
28	9/21/2018 12:56	1537531017526000	35.265096	24.087989	1566.706	161.6	-45	0	84	54	0	-45	

Appendix B. Sites selected for the study.

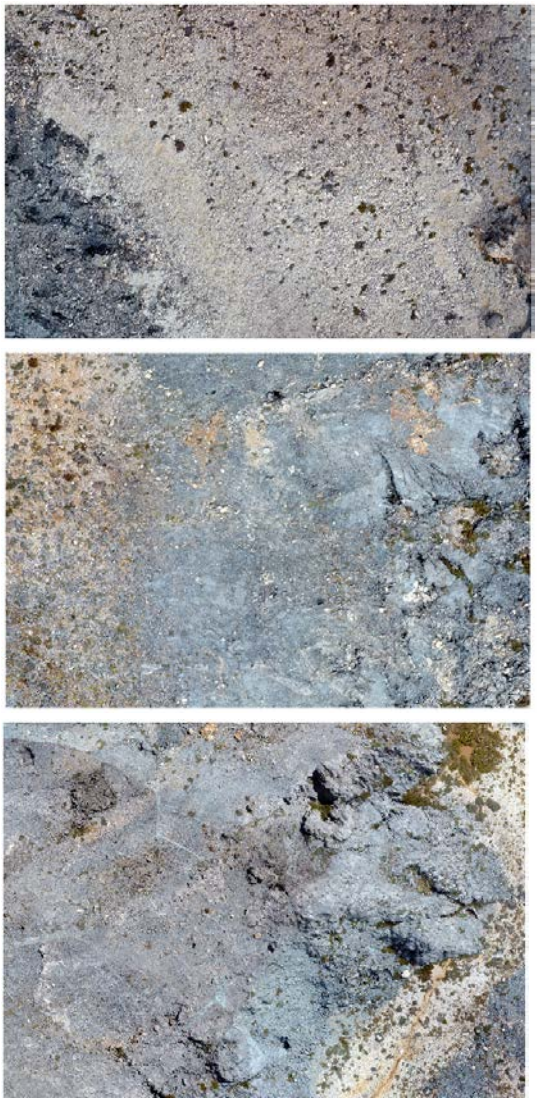
CP	Derived (Motion Image)		Measured (Orthorectified Map)		Difference	
	x	*y*	*x*	*y*	*dx*	*dy*
1	235009.3	3906565	234994	3906583	15.3	-18
2	235077.3	3906407	235072	3906415	5.3	-8
3	234989.2	3906501	234975	3906512	14.2	-11
5	235023.8	3906326	235014	3906328	9.8	-2
6	235023.1	3906436	235014	3906442	9.1	-6
8	235094.5	3906362	235095	3906366	-0.5	-4
9	235023.3	3906547	235017	3906553	6.3	-6
10	235016.7	3906353	235002	3906362	14.7	-9
11	235048.4	3906393	235036	3906400	12.4	-7
13	235095.4	3906355	235094	3906358	1.4	-3
14	235080.1	3906399	235075	3906406	5.1	-7
15	235073.8	3906435	235063	3906444	10.8	-9
16	235056.7	3906527	235048	3906531	8.7	-4
18	234887.7	3906511	234882	3906488	5.7	23
19	234896.7	3906545	234890	3906552	6.7	-7
21	234972.8	3906279	234971	3906308	1.8	-29

RMSE x= 9.19707834 RMSE y= 12.00260388 Overall RMSE = 15.12m

Appendix C. RGB images used in the study

RGB Image 1

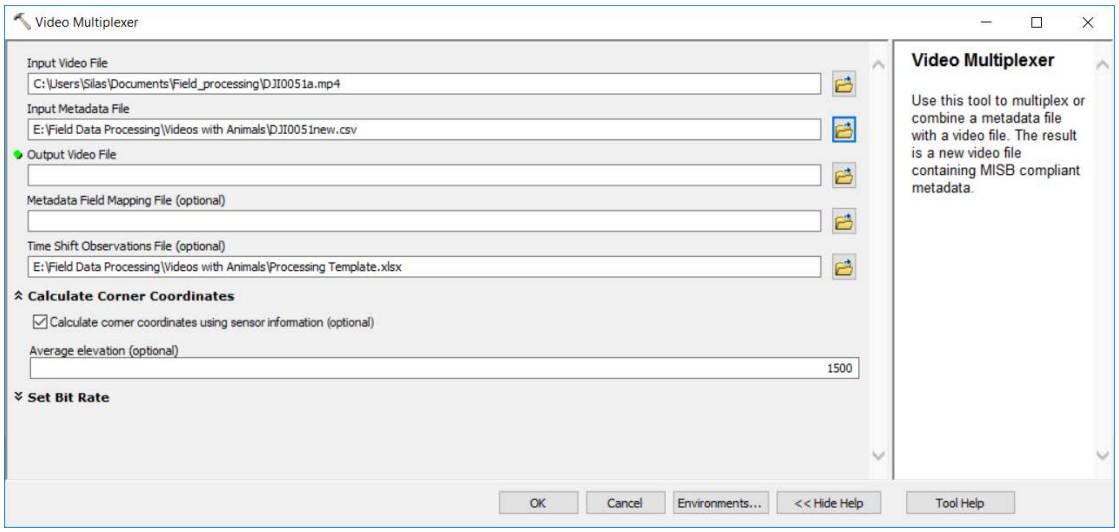
RGB Image 2



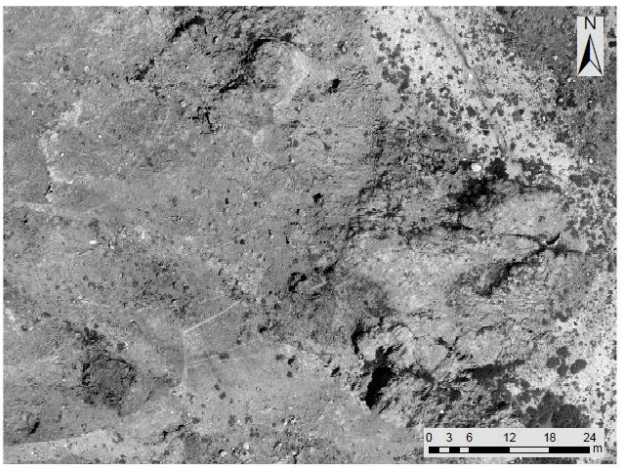

1a

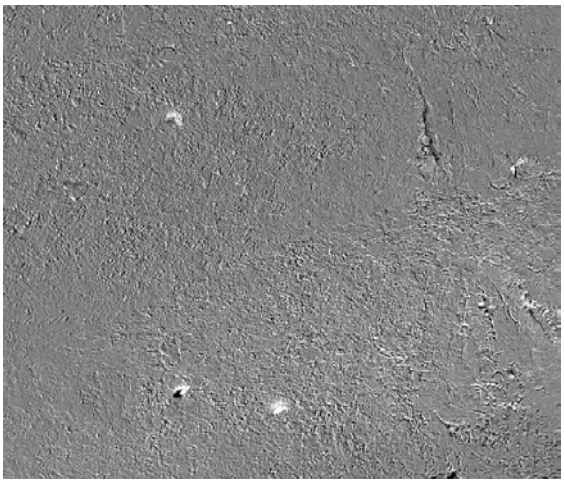

1a

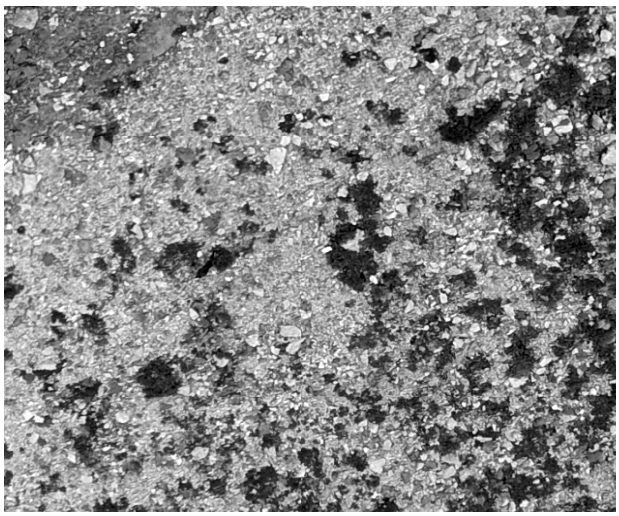
Appendix D. Video Multiplexer Interface.

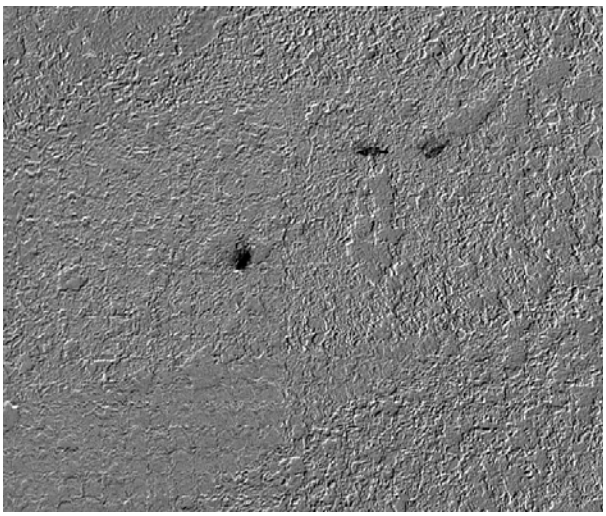


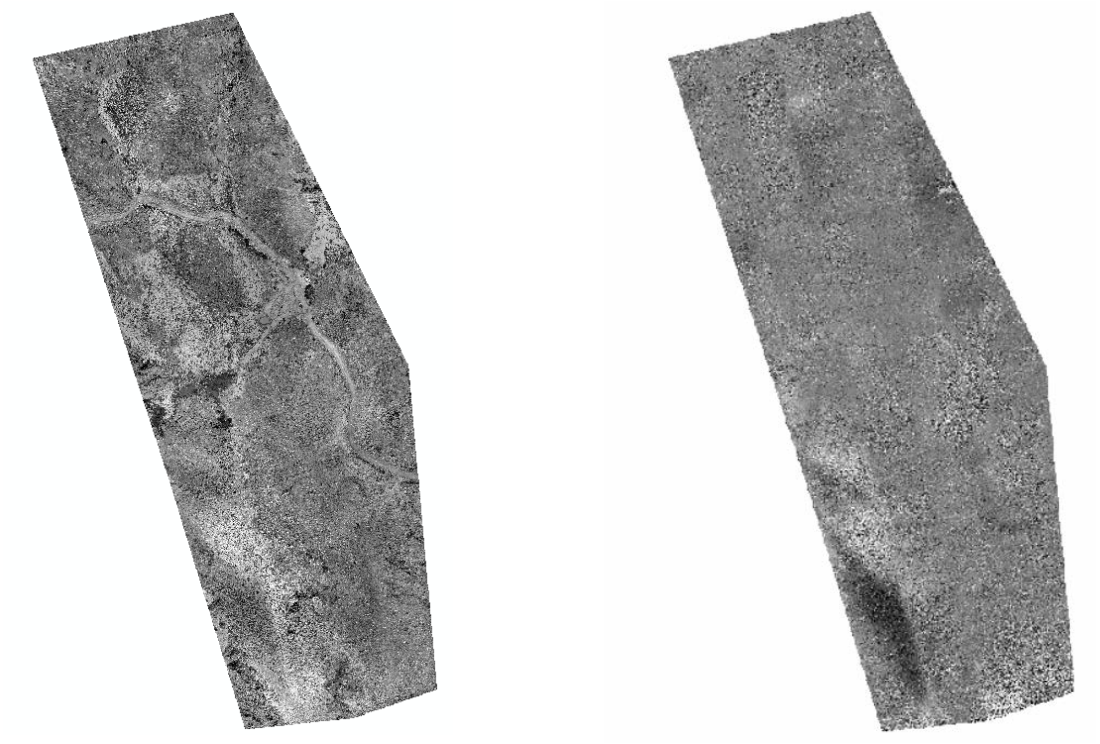
Appendix E. ICA images used in the Study



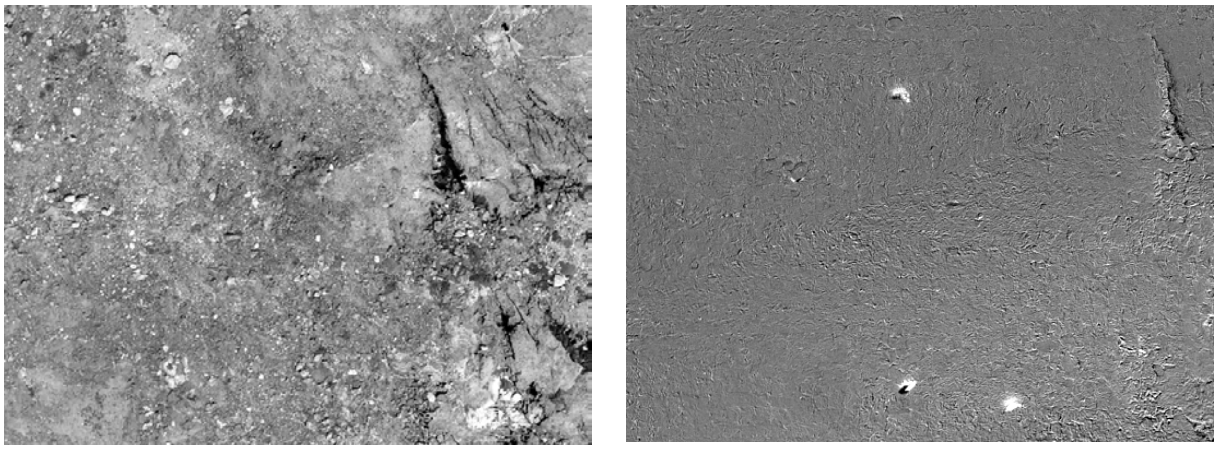


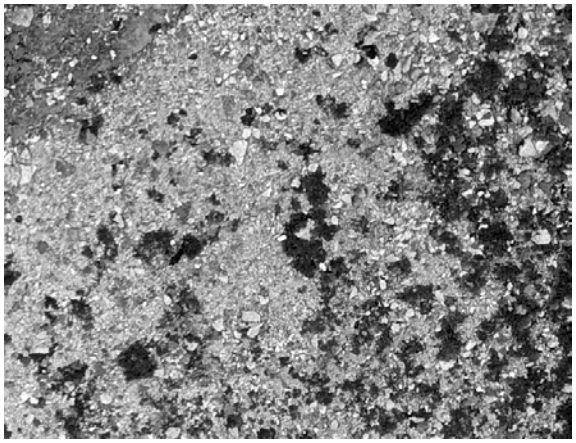


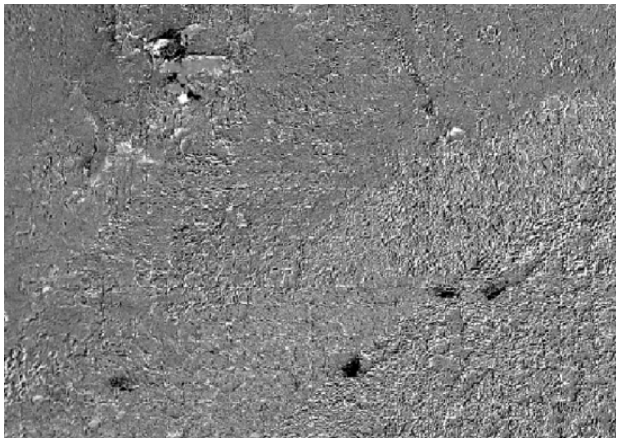


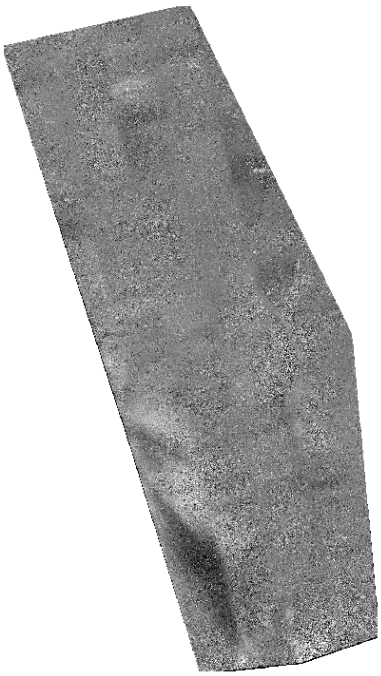


Appendix F. RGB +nDSM Fusion images









Appendix G. Field photos.