

**THE SPATIALTEMPORAL
RELATIONSHIP BETWEEN
ELEPHANT MOVEMENTS AND
THE SURFACE WATER
AVAILABILITY IN THE AMBOSELI
ECOSYSTEM**

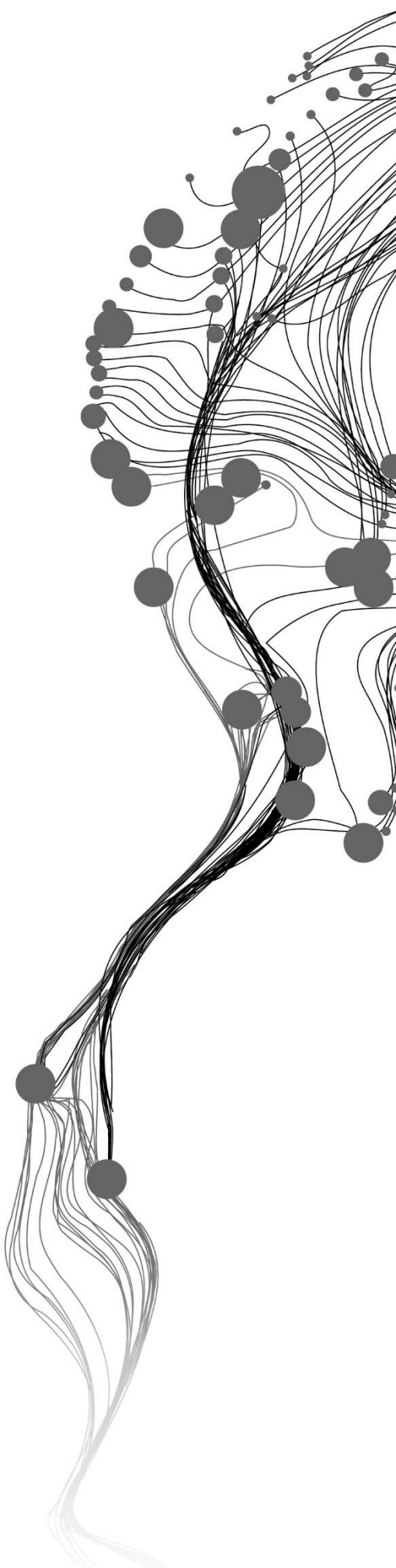
NA CHEN

February, 2019

SUPERVISORS:

Dr. Tiejun Wang

Dr. Ling Chang



THE SPATIALTEMPORAL RELATIONSHIP BETWEEN ELEPHANT MOVEMENTS AND THE SURFACE WATER AVAILABILITY IN THE AMBOSELI ECOSYSTEM

NA CHEN

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. Tiejun Wang (ITC, University of Twente)

Dr. Ling Chang (ITC, University of Twente)

ADVISORS:

Dr. Yiwen Sun (ITC, University of Twente)

Dr. Shadrack M. Ngene (Kenya Wildlife Service)

THESIS ASSESSMENT BOARD:

Prof. Dr. Andrew K. Skidmore (Chair, ITC, University of Twente)

Dr. Zoltán Vekerdy (External Examiner, ITC, University of Twente)

Dr. Tiejun Wang (First supervisor, ITC, University of Twente)

Dr. Ling Chang (Second supervisor, ITC, University of Twente)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

African elephants (*Loxodonta Africana*) are water dependent-species that can drink up to 50 gallons of water per day. Also, elephants need water for thermoregulation since they are susceptible to high temperature and sunburn. In particular, the movements of elephants are constrained by the limited surface water resources in the dry season. Despite the importance of surface water availability to elephants, lack of detailed surface water information in dry and wet season hampers further understanding of relationships between surface water and elephant movements. Remote sensing data had been widely and successfully used in the mapping of surface water. This study first mapped the surface water in the dry season and wet season over the Amboseli ecosystem used the Sentinel-1 radar image, Sentinel-2 optical image, and the combination of both based on random forest method. Then kappa coefficients of using Sentinel-1, Sentinel-2 or the combination of both were compared to identify the surface water maps that have the highest kappa coefficient for further analysis. Next, the Sentinel surface water maps were compared with existing surface water products (ILRI and JRC), after which, the distance of elephants to the nearest surface water during dry and wet season were quantified and compared. Finally, speeds of elephant movements have been calculated, classified into three categories regarding the distances of elephants to the nearest surface water and difference in the speeds was tested. These three categories are close to water (< 3km); Medium to water (3km - 6km); far from water (> 6km). Results are as follows, 1) the kappa coefficient of mapping surface water in the Amboseli ecosystem using the combination of Sentinel-1 radar and Sentinel-2 optical data is significantly higher than the one used Sentinel-1 data; 2) There is not a significant difference between the kappa coefficients of using Sentinel-2 optical data and the combination of Sentinel-1 and Sentinel-2 data. 3) A significant difference was observed in surface water distribution pattern among the three surface water products (Sentinel, ILRI and JRC). 4) The mean distance of elephants to the nearest surface water extracted from the Sentinel water products during the dry season and wet season are 6.7 km and 4.2 km, respectively; 5) The distance of elephant to the nearest surface water is longer in the dry season compared with that in the wet season; 6) A significant difference exists in the speeds of the elephant movement between close and far distance classes during the dry and wet season. The study mapped the surface water in the Amboseli ecosystem and suggested that elephant in the Amboseli ecosystem moved faster when they are closer to surface water. Findings of the relationship between the elephant movements and surface water availability are essential for understanding the movements of elephants as well as developing conservation strategies for wildlife and surface water resources.

Keywords: Surface water, elephant movement, Sentinel-1, Sentinel-2, DEM, Amboseli ecosystem, random forest

ACKNOWLEDGEMENTS

First and foremost, I want to express my deepest gratitude to my first supervisor, Dr. Tiejun Wang, who gave me tireless guidance during the entire master thesis period. This thesis would not have been able to finish without his continuous assistance. His profound knowledge and rigorous scientific attitude have influenced and will continue to inspire me throughout my lifetime.

I feel privileged to have the opportunity to be supervised by Dr. Ling Chang. I enjoyed every thought-provoking discussion with her. She is a brilliant, hard-working, considerate and promising scholar who acts as a role model that motivates me to study and strive for excellence.

Sincere acknowledgment is made to Prof. Andrew Skidmore, the chairman during my proposal and midterm presentation, who gave me critical comments and constructive suggestions that challenged me and helped me improve the thesis.

In addition, I am grateful to Dr. Yiwen Sun, Dr. Shadrack M. Ngene, Yansha Luo and Mengge Tian for providing me with valuable suggestions on the thesis, assisting me in acquiring the GPS tracking data, sharing me with the R scripts of image classification and teaching me programming, respectively.

Furthermore, I would like to thank my parents and friends, which including but not limited to Mengge Tian, Li Liu, Yayuan Lei, Shen Nie, Ruijie Li, Yan Chen, Nivedita Varma and Mst Karimon, who support me unconditionally and encourage me whenever I needed it.

Finally, my special appreciation goes to ITC for not only providing me with an international and friendly environment where I could learn GIS and remote sensing from knowledgeable professors and communicate with international classmates but also supported me with the ITC excellence scholarship. Without the scholarship, I would not have been able to start and even finish my study in the Netherlands.

Na Chen
Enschede, 25 March 2019

NOMENCLATURE

ALOS	Advanced Land Observing Satellite
AVHRR	Advanced Very High Resolution Radiometer
COSMO-SkyMed	Constellation of small satellites for the Mediterranean basin Observation
DEM	Digital Elevation Model
ERS	European Remote-Sensing Satellite
ESA	Europe Space Agency
ETM+	Enhanced Thematic Mapper Plus
GLCM	Gray Level Co-Occurrence Matrix
GPS	Global Positioning System
GRD	Ground Range Detected
ILRI	International Livestock Research Institute
JRC	Joint Research Centre
MDG	Mean Decrease Gini
MODIS	Moderate Resolution Imaging Spectroradiometer
MRF	Markov Random Field
MSAVI	Modified Soil Adjusted Vegetation Index
MSS	Multispectral Scanner Sensor
MSI	Multispectral Instrument
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIMA	National Imagery and Mapping Agency
OLI	Operational Land Imager
OOB	Out of bag
Radar	Radio Detection And Ranging
Radarsat	Canada's Radar Earth Observation Satellite
RCMRD	Reginal Centre for Mapping of Resources for Development Geoportal
SAR	Synthetic Aperture Radar
SLC	Single Look Complex
SNAP	Sentinel Application Platform
SRTM	Shuttle Radar Topography Mission
TanDEM-X	TerraSAR-X add-on for Digital Elevation Measurements
TerraSAR-X	German Radar Earth Observation Satellite
TM	Thematic Mapper
UTM	Universal Transverse Mercator

TABLE OF CONTENTS

1.	Introduction	1
1.1.	Background.....	1
1.2.	Problem statement	3
1.3.	Research objectives	4
1.4.	Research questions	4
1.5.	Research hypotheses	4
1.6.	Outline of the thesis.....	5
2.	Study area and data.....	6
2.1.	Study area.....	6
2.2.	Sentinel-1 radar data.....	7
2.3.	Sentinel-2 optical data.....	8
2.4.	Elephant GPS tracking data	9
2.5.	Ancillary data.....	10
3.	Methodology	12
3.1.	Data processing.....	12
3.2.	Mapping surface water using random forest	15
3.3.	Comparing surface water products	21
3.4.	Quantifying and comparing the distance of elephants to the nearest surface water	22
3.5.	Analyzing the relationship between elephant movement and the surface water availability	22
4.	Results	24
4.1.	Surface water maps	24
4.2.	Comparison of the surface water.....	31
4.3.	The distance of elephants to the nearest surface water	34
4.4.	The relationship between speeds of elephant and the surface water availability	35
5.	Discussion.....	37
5.1.	Remote sensing of mapping surface water in the Amboseli ecosystem	37
5.2.	The relationship between the surface water availability and elephant movement in the Amboseli ecosystem.....	38
6.	Conclusions and Recommendations	40
6.1.	Conclusions	40
6.2.	Limitations and suggestions for future research.....	41

LIST OF FIGURES

Figure 1 Study area and the four collared elephants in 2015 in the Amboseli ecosystem.....	6
Figure 2 Monthly average rainfall in the Amboseli ecosystem (2010-2017), Source: World Weather.....	7
Figure 3 Sentinel-1 images with VV polarizations used in this study. (A) is the image acquired on July 6, 2016,.....	7
Figure 4 Sentinel-1 images with VH polarizations used in this study. (A) is the image acquired on July 6, 2016,.....	8
Figure 5 True color composite (R, G, B =Band4, 3, 2) of Sentinel-2 images used in this study. (A) is the image acquired on July 23, 2016. (B) is the image acquired on December 20, 2016.....	8
Figure 6 GPS fixes of four collared elephants in 2015 displayed in UTM Zone 37S coordinate system. KM, KUF, MAM and ELM are names of elephants.....	9
Figure 7 True color composite of Planet scope image (a subset of the study area) acquired on July 24, 2016.....	10
Figure 8 Surface water map in the Amboseli ecosystem provided by ILRI.....	11
Figure 9 Surface water maps in dry season (left) and wet season (right) over the Amboseli ecosystem provided by JRC.....	11
Figure 10 True color composite (R, G, B =Band4, 3, 2) of Sentinel-2 image acquired on July 23, 2016 (A), and the corresponding cloud and shadow mask generated from Fmask 4.0 software (B).....	14
Figure 11 True color composite (R, G, B =Band4, 3, 2) of Sentinel-2 image acquired on December 20, 2016 (A) and the corresponding cloud and shadow mask generated from Fmask 4.0 software (B).	14
Figure 12 Distribution of training and test points in the Amboseli ecosystem during the dry season (A) and wet season (B).....	16
Figure 13 Variable importance of random forest classification results using different datasets.....	17
Figure 14 Tuning of “mtry” parameter of six different random forest models using different imagery.	19
Figure 15 Effects of number of trees on the OOB error rate based on the training data of six random forest models. Black line represents the overall OOB error rate of all the classes in random forest classification. While green and red lines represent the OOB error rate of water and Non-water class in the random forest models. The horizontal axis represents the parameter “ntree”, number of trees while the y axis represents the OOB error rate.....	20
Figure 16 Surface water map of the Amboseli ecosystem during the dry season of 2016 based on Sentinel-1 radar imagery and DEM data.....	24
Figure 17 Surface water map of the Amboseli ecosystem during the wet season of 2016 based on Sentinel-1 radar imagery and DEM data.....	25
Figure 18 Surface water map of the Amboseli ecosystem during the dry season of 2016 based on Sentinel-2 optical imagery and DEM data.....	26
Figure 19 Surface water map of the Amboseli ecosystem during the wet season of 2016 based on Sentinel-2 optical imagery and DEM data.....	27
Figure 20 Surface water map of the Amboseli ecosystem during the dry season of 2015 based on Sentinel-1 radar imagery, Sentinel-2 optical imagery and DEM data.....	28
Figure 21 Surface water map of the Amboseli ecosystem during the wet season of 2016 based on Sentinel-1 radar imagery, Sentinel-2 optical imagery and DEM data.....	29
Figure 22 Comparison of kappa coefficients during the dry season (A) and wet season (B).	30
Figure 23 Euclidean distance to surface water in dry and wet season calculated from Sentinel water product.....	31

Figure 24 Euclidean distance to surface water in dry and wet season calculated from ILRI water product.32

Figure 25 Euclidean distance to surface water in dry and wet season calculated from JRC water product. 32

Figure 26 Distances of 3000 random points to the surface water during the dry season (A) and wet season (B). The vertical axis is the distance (km).....33

Figure 27 Distance of elephants to the nearest surface water in the Amboseli ecosystem during the dry and wet seasons using Sentinel (A), JRC (B), and ILRI (C) surface water product. The pairs with no significant difference (C) based on the Mann-Whitney U test with $p > 0.05$ are annotated.....34

Figure 28 Comparison of speed of elephant movement among three distance to water categories based on surface water derived from Sentinel satellite images. (A) represent the difference in speeds of elephant movement during the dry and (B) represent the one in wet season. The pairs with no significant difference based on the Mann-Whitney U test ($p > 0.05$) are annotated.....35

Figure 29 Speeds of elephant movement among three distance to water categories based on surface water produced by ILRI during the dry season(A) and wet season (B). The pairs with no significant difference are annotated based on Mann-Whitney U test with $p > 0.05$36

LIST OF TABLES

Table 1 Basic information about the four collared elephants in the Amboseli ecosystem.....	9
Table 2 Formulas of indices in Sentinel-1 radar data.....	13
Table 3 Formulas of water and vegetation indices in Sentinel-2 satellite images (Green = Band 3, Red = band 4, near infrared (NIR)= Band 8).	13
Table 4 Input variables of random forest models using different datasets.	16
Table 5 Tuned parameters for random forest models used to map surface water in the Amboseli ecosystem.	18
Table 6 Descriptive statistics of elephant movement speed (Unit: km/h) during the dry and wet seasons.	23
Table 7 The confusion matrix of classification using Sentinel-1 radar data and DEM during the dry season.	24
Table 8 The confusion matrix of classification using Sentinel-1 radar and DEM data during the wet season.	25
Table 9 The confusion matrix of classification using Sentinel-2 optical data and DEM during the dry season.	26
Table 10 The confusion matrix of classification using Sentinel-2 optical and DEM data during the wet season.	27
Table 11 The confusion matrix of classification using Sentinel-2 optical and DEM data during the wet season.	28
Table 12 The confusion matrix of classification using Sentinel-1, Sentinel-2 and DEM during the wet season.	29
Table 13 Comparison of the number and area of surface water.....	31

1. INTRODUCTION

1.1. Background

Water is of great importance to all living things (Al-weshah, 2016). For example, water is an integral part of animal organisms, which accounts for 65% to 70% of animals body weights (Ganguly & Delhi, 2016). Water is vital for regulating the processes of digestion, transportation of nutrients, excretion and body temperature of wild animals (Smathers, 2018). Animals will die in a few days without drinking water (Popkin & Rosenberg, 2011). In addition, water provides habitats for countless animal species, including fish, amphibians, reptiles, insects, birds, and mammals, in the form of rivers, lakes, swamps or ocean (Klapproth & Johnson, 2009). Some animals, mainly aquatic species are entirely dependent on water, they live in water for their whole life. While other animals frequently visit water to drink and use water for water-related activities, like mudding, bathing, and swimming. Furthermore, the presence of water can affect the distribution of wild animals. Previous studies have shown that for some species, e.g., large herbivore, the distribution of which is influenced by the location of water, especially during the dry season (Bergstrom & Skarpe, 1999; Redfern, Grant, Biggs, & Getz, 2015; Western, 1975). Due to the importance of water to animals, and the fact that water is unevenly distributed in space and time (Yano, Hanasaki, Itsubo, & Oki, 2015), animals have to move in heterogeneous landscapes to search for water to increase their chances for survival (de Beer & van Aarde, 2008).

Elephants, as the largest terrestrial herbivores, can drink up to 50 gallons (about 200 liters) water within one day (Dirzo et al., 2015). Also, elephants are observed to be sensitive to high temperature and sunburn (Fuller, Mitchell, Maloney, & Hetem, 2016). Thus, elephants need a large amount of water for activities related to water, like mud bathing, swimming for regulating body temperature (Dunkin, Wilson, Way, Johnson, & Williams, 2013; Purdon & van Aarde, 2017). African elephants (*Loxodonta Africana*) are typical water-dependent species (Gara et al., 2017; Makindi, Mutinda, Olekaikai, Olelebo, & Abdillahi, 2014; Western, 1975). However, the surface water (including rivers, lakes, and marshes) in the semi-arid landscape is highly seasonal, which makes the distributions and movements of elephants constrained by the surface water availability in the dry season. Western (1975) studied the influence of water availability on the structure and seasonal dynamic changes of large mammals in the Amboseli ecosystem of Kenya and found that water-dependent species, including elephants, are concentrated near permanent water bodies during dry seasons, and they disperse as water availability increases during the wet season. Chamaillé-Jammes, Valeix, and Fritz (2007) investigated how elephant density and surface water availability influenced each other and concluded that surface water drives the distribution and abundance of elephants within Hwange National Park, Zimbabwe. de Beer et al. (2008) examined the relationship between elephants home range and water distribution and concluded that the home range size of elephants decreases with increasing water point density. Chamaillé-Jammes et al. (2013) studied the movement speed of elephants and found that elephants increase their speed when they are close to water points. Redfern et al. (2015) investigated the elephant movement pattern and pointed out that surface water constrains the pattern of elephant movement. In a word, surface water availability and its spatial distribution are important for the distribution and movements of elephants in a semi-arid landscape, especially in dry season. In the semi-arid landscape, surface water becomes scarce due to little precipitation in dry seasons (Chamaillé-Jammes et al., 2013), elephants even take the risk of being expelled by pastoralists, battle for water with livestock around the artificial water points. Droughts are becoming more and more frequent in arid and semi-arid landscapes across Africa over the last few decades (Easterling et al., 2000). Once this extreme weather event happens, surface water

becomes increasingly scarce, natural mortality of elephants, especially juvenile elephants will increase because elephants will perish of thirst if they travel too far to search for water. For example, Ngene et al., (2013) have studied the elephant population in the Tsavo-Makomazi ecosystem, Kenya and observed that in climatically normal years, e.g., 2005, 2006, 2007 and 2008, number of dead elephants caused by natural factors is less than 30. However, this number increased to about 400 in 2009, when a severe drought occurred. Despite the importance of surface water availability to elephants, the lack of detailed surface water information across space and time hampers further understanding of relationships between surface water and elephant movements.

Satellite images have been extensively and increasingly adopted to map surface water due to its low-cost, large spatial coverage, efficiency advantages over in situ measurements (Clement, Kilsby, & Moore, 2017). According to the source of energy, two types of satellite sensors, active and passive sensors, can be used to acquire satellite images. The active sensor uses its own source of energy to illuminate objects, while passive sensor takes advantage of the energy emitted by the sun and then reflected by objects. Optical satellite images are acquired from the passive sensor, while the Synthetic Aperture Radar (SAR) imagery is obtained from the active sensor.

The capability of optical satellite images for surface water mapping and monitoring has been well demonstrated since the 1970s. Optical satellite images at different spatial resolution, coarse, medium or high spatial resolution, acquired from various sensors can and have been used for water detection (Huang, Chen, Zhang, & Wu, 2018). Extensive research has been conducted to detect surface water using coarse and medium resolution imagery acquired from NOAA/AVHRR (Dietz et al., 2017), MODIS (Li, Sun, Goldberg, & Stefanidis, 2013), Landsat series (MSS, TM, ETM+, OLI) (Du et al., 2016; Kaplan & Avdan, 2017) and Sentinel-2 MSI. Similarly, the high spatial resolution satellite images are also applicable in monitoring surface water extent. For example, the Worldview, RapidEye, Qucikbird, GF-1, GF-2, and ZY-3 images also have been utilized to detect surface water in previous studies and results showed that high spatial resolution satellite images have advantages for detecting small water bodies (Nath & Deb, 2010). All kinds of satellite imagery can successfully detect surface water as the water has its low spectral reflectance in infrared bands (Huang et al., 2018). But there is a trade-off among the spatial, spectral and temporal resolution when selecting suitable images for surface water mapping. No matter what kind of optical satellite images have been selected, a fact that will never change is that clouds can affect optical images very easily, especially in frequently cloudy and rainy areas, which limits its potential for spatial-temporal monitoring of surface water dynamics.

SAR sensor has considerable advantages compared with the optical sensor for quantifying spatial-temporal variations of surface water (Huang et al. 2018). SAR not only has the capability of penetrating clouds but also being sensitive to open water and below-canopy water (Kim et al., 2014; Martinis et al., 2015a). Since smooth water appears dark in radar images due to specular reflection, the potential of detecting and monitoring surface water using SAR data has been well examined and demonstrated in literature by using X-band (COSMOSkyMed, TerraSAR-X, TanDEM-X, SIR-C), C-band (ERS-1, ERS-2, Sentinel-1, RADARSAT-1/2, SIR-C) and L-band (SEASET, SIR-A, SIR-B, JERS-1, ALOS-1, ALOS-2). For instance, Katherrine et al. (2018) compared single-polarization and dual-polarisation in surface water monitoring using TerraSAR-X and the study observed that the dual-polarisation has advantages over single-polarization for surface water mapping. Aldenhoff, Heuzé, and Eriksson (2018) compared C-band (Sentinel-1) and L-band (ALOS-2) for mapping water using a Neural Network algorithm, and results showed that classification accuracies derived from C-band and L-band are 95%, 88%, respectively. Despite the potential of detecting surface water using SAR imagery, the technical challenges still exists. Water-alike objects, like shadows, sands, and pavements usually cause inaccurate estimation of surface water (Martinis et al., 2015b). Also,

unexpected changes in the backscatter of surface water can be resulted from wind-induced Bragg scattering and volume scattering due to inundated vegetation (Bragg et al., 2007).

1.2. Problem statement

Elephants are water-dependent animals, and the elephant is hypothesized to stay close to the surface water. However, the distances of the elephant to the surface water has not reached an agreement in previous studies. Also, only a few studies explored whether elephants stay closer to the surface water in the dry season compared with wet season. De Beer et al. (2006) used the artificial and natural water points, and 17-years GPS collared data of six elephants to study the interaction between elephants, plants, and water, they concluded that elephants preferred to stay in the area within 4 km from water sources. Similarly, elephants in Chobe National Park, Botswana were observed to aggregate within 3.5 km from surface water in the dry season and 5 km from surface water in the wet season (Park, Stokke, & Toit, 2002). Also, elephants in Etosha National park and Tembe Elephant Park of southern Africa were observed to stay closer to water during the dry season compared with wet season (Harris, Russell, Van Aarde, & Pimm, 2008). Western (1975) studied the effect of water availability on the elephants distribution and found that elephants concentrate in the area within 2 km, 6 km to 8 km from water. However, Viljoen (1989) pointed out that elephants in South West Africa regularly travel 20 km to 40 km without drinking water. Studies also has been carried out to analyze the relationship between elephant occurrence and water sources, and the result shows that elephant occurrence can reach peak values even if the distance between elephant and water resources ranges from 20 km to 60 km during the dry season (Ndaimani, Murwira, Masocha, & Zengeya, 2017).

Previous researches either use the permanent water maps or artificial water points, instead of using dynamic water maps in dry and wet seasons to study the relationship between surface water availability and elephant movements. For a further understanding of the relationships between elephant movements and surface water in dry and wet seasons, surface water maps in different seasons extracted from satellite images with acceptable spatial resolution are needed.

The advantages of mapping surface water using either optical satellite images or SAR imagery have been well demonstrated in (Irwin, Braun, Fotopoulos, Roth, & Wessel, 2018; Kaplan & Avdan, 2017; Nath & Deb, 2010). Despite the advantages of mapping surface water using either optical satellite images or SAR imagery, the disadvantages of mapping surface water using either Sentinel-1 or Sentinel-2 imagery cannot be ignored. For optical images, the cloud contamination issue of optical satellite images is unavoidable even adopting well-defined algorithms (Ogunbadewa, 2012). For SAR imagery, the water-like objects will introduce errors in the surface water map due to the occurrence of specular reflection. To remedy these limitations, a few studies have been conducted to combine optical and SAR imagery to map surface water (Markert, Chishtie, Anderson, Saah, & Griffin, 2018). Whether combining optical and SAR could significantly increase the classification accuracies in surface water mapping remains unknown. On one hand, the optical and SAR imagery contain complementary information (Amarsaikhana et al., 2010). On the other hand, errors resulted from each dataset can cause error propagation in the final maps (Liu, Gopal, & Woodcock, 2004; Markert et al., 2018). Only a few studies have been carried out to map surface water using combined Sentinel-1 SAR and Sentinel-2 optical data, not to mention the comparison of the accuracies of maps produced by individually using of Sentinel-1, Sentinel-2 or the combined Sentinel-1 and Sentinel-2 data.

1.3. Research objectives

The main objective of the study is to examine the spatiotemporal relationships between elephant movements and the surface water availability derived from satellite remote sensing in the Amboseli ecosystem, Kenya. Specific research objectives are as follows:

- To map the surface water in the Amboseli ecosystem during the dry and wet seasons using Sentinel-1 radar data, Sentinel-2 optical data and the combination of Sentinel-1 and Sentinel-2 data, respectively.
- To compare the surface water maps derived from Sentinel images with the other two existing surface water products produced by the Joint Research Center (JRC) and the International Livestock Research Institute (ILRI), respectively.
- To quantify and compare the distance of elephants to the nearest surface water in the Amboseli ecosystem between the dry season and wet season.
- To examine the relationship between the movement speeds of elephants and the distance to surface water in the Amboseli ecosystem during the dry and wet seasons, respectively.

1.4. Research questions

- Does the combined Sentinel-1 radar data and Sentinel-2 optical data significantly improve the surface water mapping accuracy?
- Is there a statistically significant difference in surface water distribution pattern (e.g., number of water bodies or rivers, mean distance to surface water) among three surface water products (i.e., Sentinel, JRC, and ILRI)?
- Is there a statistically significant difference between the distance of elephant to the nearest surface water during the dry and wet seasons?
- What is the relationship between the movement speed of elephants and the distance to surface water in the Amboseli ecosystem during the dry season and wet season, respectively?

1.5. Research hypotheses

Hypothesis 1

H_0 : There is no statistically significant difference in surface water mapping accuracies between the use of the Sentinel-1 (or Sentinel-2) data and the combination of Sentinel-1 and Sentinel-2 data.

H_1 : The surface water mapping accuracy derived from the combined Sentinel-1 and Sentinel-2 data is significantly higher than the accuracies derived from either Sentinel-1 or Sentinel-2 alone.

Hypothesis 2

H_0 : There is not any statistically significant difference in surface water distribution pattern among the three surface water products.

H_1 : A statistically significant difference exists in surface water distribution pattern among the three surface water products.

Hypothesis 3

H_0 : There is no statistically significant difference in the distance of elephants to the nearest surface water between the dry and wet seasons in the Amboseli ecosystem.

H_1 : The distance of elephants to the nearest surface water in the dry season is significantly larger than that of the wet season in the Amboseli ecosystem.

Hypothesis 4

H0: There is no significant difference in the movement speeds of elephant whether the elephants are closer or further away from the surface water during the dry and wet seasons, respectively.

H1: The movement speeds of elephant are higher when they are close to surface water compared with further away from surface water during the dry season and wet season, respectively.

1.6. Outline of the thesis

The thesis consists of six chapters. More specifically,

Chapter 1 introduction describes the background, problem statement, research objectives, research questions, research hypothesis and outline of the thesis.

Chapter 2 study area and data consist of study area and data. The study area mainly illustrates the geographic location, the climate characteristics and area of the Amboseli ecosystem. The data part describes the sources and characteristics of Sentinel-1 radar data, Sentinel-2 optical data, ancillary data and elephant GPS tracking data.

Chapter 3 methodology reviews the methods used and illustrates the reasons for adopting a particular method and details about the procedure of data processing. This part mainly explains the surface water mapping using random forest algorithm and the method to analyze the relationship between the dynamic distribution of surface water and the elephant movements.

Chapter 4 results present the findings of the random forest image classification and the relationship between the elephant movements and surface water availability.

Chapter 5 discussion demonstrates the significance, limitations and potential improvements in the future work of this study. In addition, this part examines the interpretation of results from different perspectives.

Chapter 6 conclusions and recommendations. The conclusions answered the research questions, proposed in the introduction section, based on the results. The recommendation summarizes the limitations of this study and proposes suggestions for future research.

2. STUDY AREA AND DATA

This chapter consists of five sections. Section 2.1 mainly illustrates the geographic location, the climate type and characteristics of the study area, the Amboseli ecosystem. The following four sections mainly describe the sources and the characteristics of different data used in this study. Section 2.2 elaborates the acquisition and specifications of Sentinel-1 radar imagery. Section 2.3 describes the source and explains the source, wavelengths intervals and spatial resolution of multispectral bands in Sentinel-2 imagery. Section 2.4 demonstrates the source, acquisition and detailed information about the ID, sex, name and number of total and used GPS fixes in this study. Section 2.5 illustrates the ancillary data, and this section describes the sources and characteristics of Planet Scope image, DEM, surface water products provided by International Livestock Research Institute (ILRI) and Joint Research Center (JRC) of Europe Commission.

2.1. Study area

The Amboseli ecosystem locates in Kajiado County, Kenya (Gara et al., 2017) (Figure 1). This ecosystem mainly consists of the Amboseli National Park and surrounding rangelands (Mose et al., 2013; Shadrack et al., 2017) and the area of the study area is approximately 11000 km². A series of swamps, the area where water is partially or completely submerged by vegetation, refilled by mountain forest catchment in this ecosystem supports a large number of mammals (Mose et al., 2013). The climate type of this ecosystem is Savanna climate, also known as tropical wet and dry climate (Köppen, 1884), the typical characteristics of which are high temperature all year round and with the distinct dry and wet season. In this ecosystem, the average temperature of each month ranges from 20 °C to 30 °C (Gara et al., 2017), the annual precipitation of this ecosystem is spatially and temporally heterogeneous, ranges from 200 mm to 1000 mm (BurnSilver, Worden, & Boone, 2008). The dry period falls in June, July, August, and September, other months have relatively higher average rainfalls (Figure 2). Surface water is limited across the ecosystem and permanent water mainly consists of rivers and a series of swamps (BurnSilver et al., 2008). The estimated population of elephants in the Amboseli ecosystem is approximately 1500 (Okello et al., 2016). Historically, elephants concentrate on the swamps in the dry season and disperse in the ecosystem during the wet season (Western, 1975).

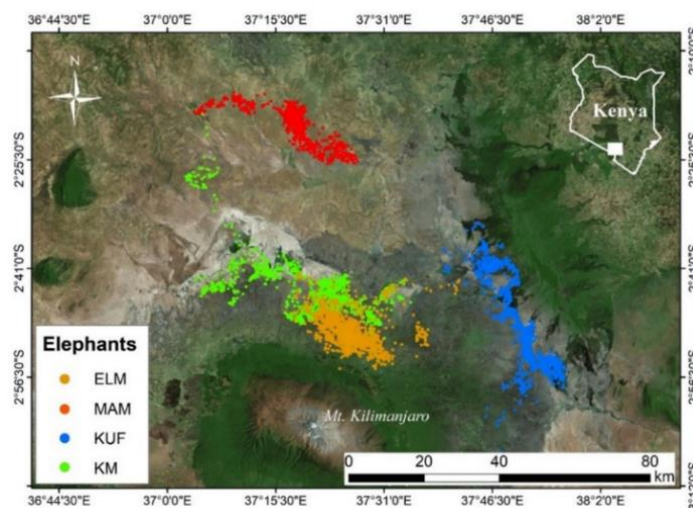


Figure 1 Study area and the four collared elephants in 2015 in the Amboseli ecosystem.

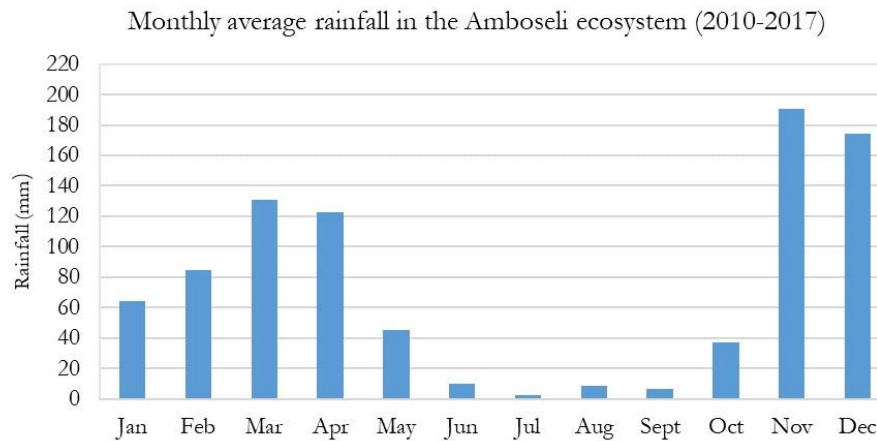


Figure 2 Monthly average rainfall in the Amboseli ecosystem (2010-2017), Source: World Weather.

2.2. Sentinel-1 radar data

Sentinel-1 is designed for the Copernicus program (earth observation program) and developed by the European Space Agency. Sentinel-1 is a radar mission that consists of Sentinel-1A and Sentinel-1B satellites, launched in France Guiana spaceport on April 3, 2014 and April 25, 2016, respectively (Martín Serrano et al., 2015). Sentinel-1 is equipped with a 5.405 GHz C-band SAR sensor that can provide day and night, all-weather imagery services with short revisit time (Aschbacher & Milagro-Pérez, 2012). Sentinel-1 imagery can be acquired through various modes with different spatial resolutions and swath widths (Huang et al., 2018). IW mode (Interferometric Wide Swath) data is acquired at the primary operational mode over land and the most widely used data with a swath width of 250 km at 5 m * 20 m (range and azimuth) spatial resolution (Huang et al., 2018; Zeng et al., 2017). IW mode can produce data at different levels, namely Level-0, Level-1, and Level-2. Level-0 is raw data. Level-1 is designed for most users and consists of Single Look Complex (SLC) and Ground Range Detection (GRD) products. The SLC products include phase information, while the GRD products do not. Also, GRD products are multi-looked with reduced speckle as well as reduced spatial resolution (20 * 22 m) with squared pixels (Tapete & Cigna, 2017). Level-2 products are designed for ocean monitoring, contain information about wind speed and direction. Sentinel-1 is a dual polarization system that can transmit and receive in Horizontal and Vertical direction. Figure 3 and Figure 4 show the Sentinel-1 imagery with VV polarizations and VH polarizations used in this study.

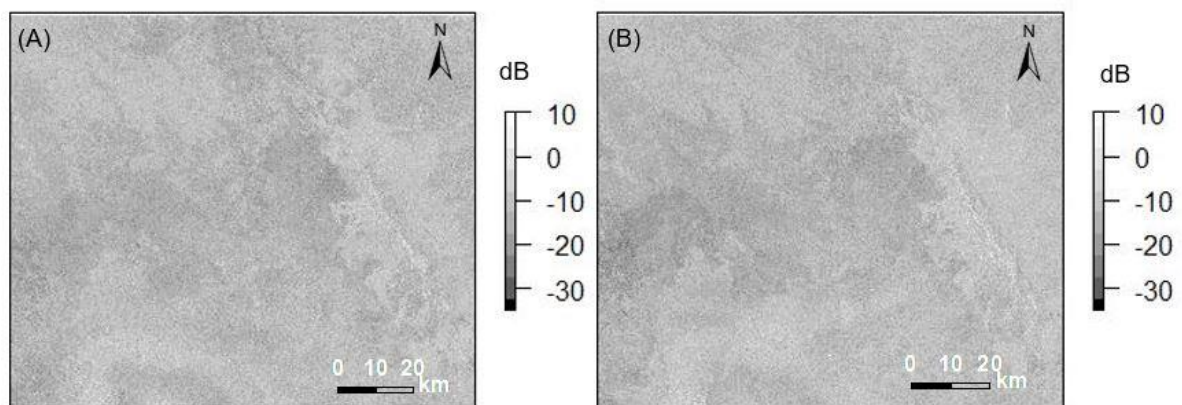


Figure 3 Sentinel-1 images with VV polarizations used in this study. (A) is the image acquired on July 6, 2016, (B) is the image mosaicked from December 9, 2016 and December 14, 2016.

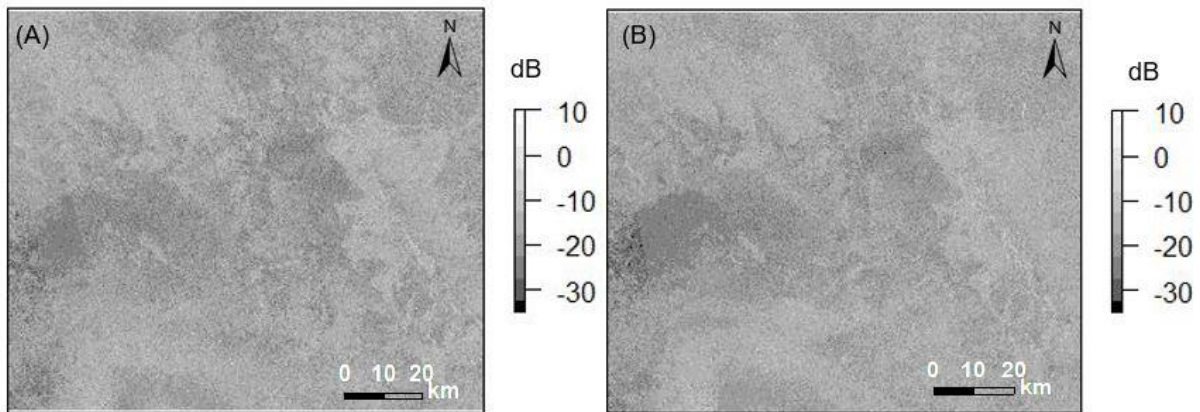


Figure 4 Sentinel-1 images with VH polarizations used in this study. (A) is the image acquired on July 6, 2016, (B) is the image mosaicked from December 9, 2016 and December 14, 2016.

2.3. Sentinel-2 optical data

Sentinel-2 is a multi-spectral high-resolution mission consists of Sentinel-2A and Sentinel-2B satellites, launched on June 23, 2015 and March 7, 2017, respectively (Li & Roy, 2017). A Sentinel-2 image consists of 13 spectral bands (wavelength ranges from 443 nm to 2190 nm), the spatial resolution ranges from 10 m to 60 m. Band 2 (490 nm), Band 3 (560 nm), Band 4 (665 nm) and Band 8 (842 nm) are at 10 m spatial resolution. 60 m resolution bands consist of 4 vegetation red edge bands (705 nm, 740 nm, 783 nm, 865 nm) and 2 short wave infrared (SWIR) bands (1610 nm, 2190 nm) designed for snow, ice or cloud detection (Aschbacher & Milagro-Pérez, 2012). The remained three bands (443 nm, 945 nm, 1375 nm) designed for cloud masking and atmospheric correction are at 60 m resolution (Aschbacher & Milagro-Pérez, 2012). Figure 5 shows the used Sentinel-2 images in this study.

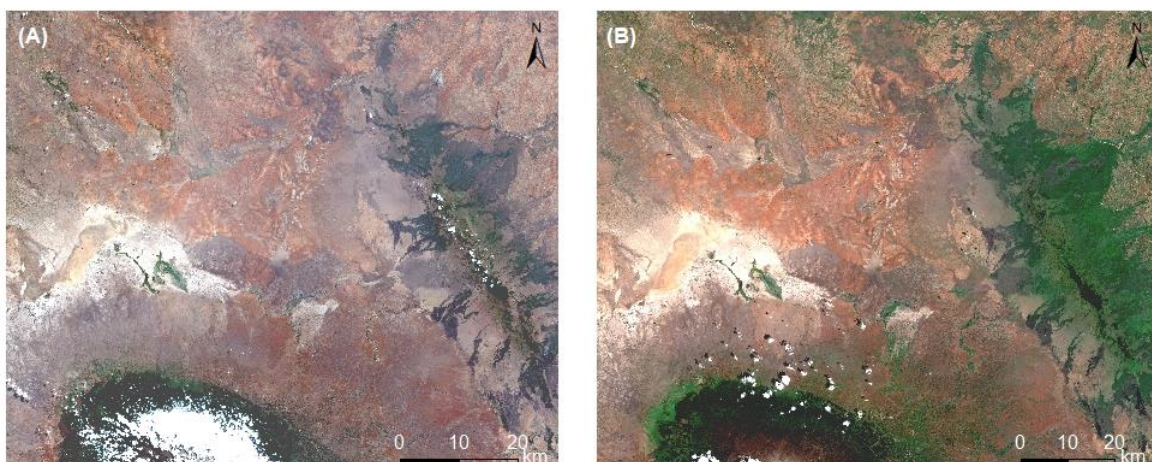


Figure 5 True color composite (R, G, B =Band4, 3, 2) of Sentinel-2 images used in this study. (A) is the image acquired on July 23, 2016. (B) is the image acquired on December 20, 2016.

2.4. Elephant GPS tracking data

Sponsored by the International Fund for Animal welfare (IFAW), Kenya Wildlife Services (KWS) initiated the collaring operation for elephants in the Amboseli ecosystem on 18 February, 2013. Each elephant was injected 18 mg Etorphine hydrochloride with Dan-Inject gun and revived by injecting 54mg diprenorphine at the ear vein. The GPS collar captured one fix (a point with geographic coordinates) every four hours using GMT +3 time (three hours ahead of Greenwich Mean Time). The detailed information of four collared elephants, such as the ID, name, sex, group size, tracking period, number of the total and used GPS fixes of the collared elephants are described in Table 1. Each collared elephant represents one family group. Figure 6 shows the GPS fixes of the four collared elephants in 2015.

Table 1 Basic information about the four collared elephants in the Amboseli ecosystem.

ID	Name of elephants	Sex	Age	Group size	Tracking period	Number of GPS fixes	Used GPS fixes
1	KM	Male	26	5	02/19/2013-04/06/2016	6254	678
2	KUF	Female	26	10	03/14/2013-04/06/2016	6031	670
3	MAM	Male	25	3	12/04/2013-04/06/2016	4171	616
4	ELM	Male	20	Unknown	04/23/2014-12/17/2015	3258	600

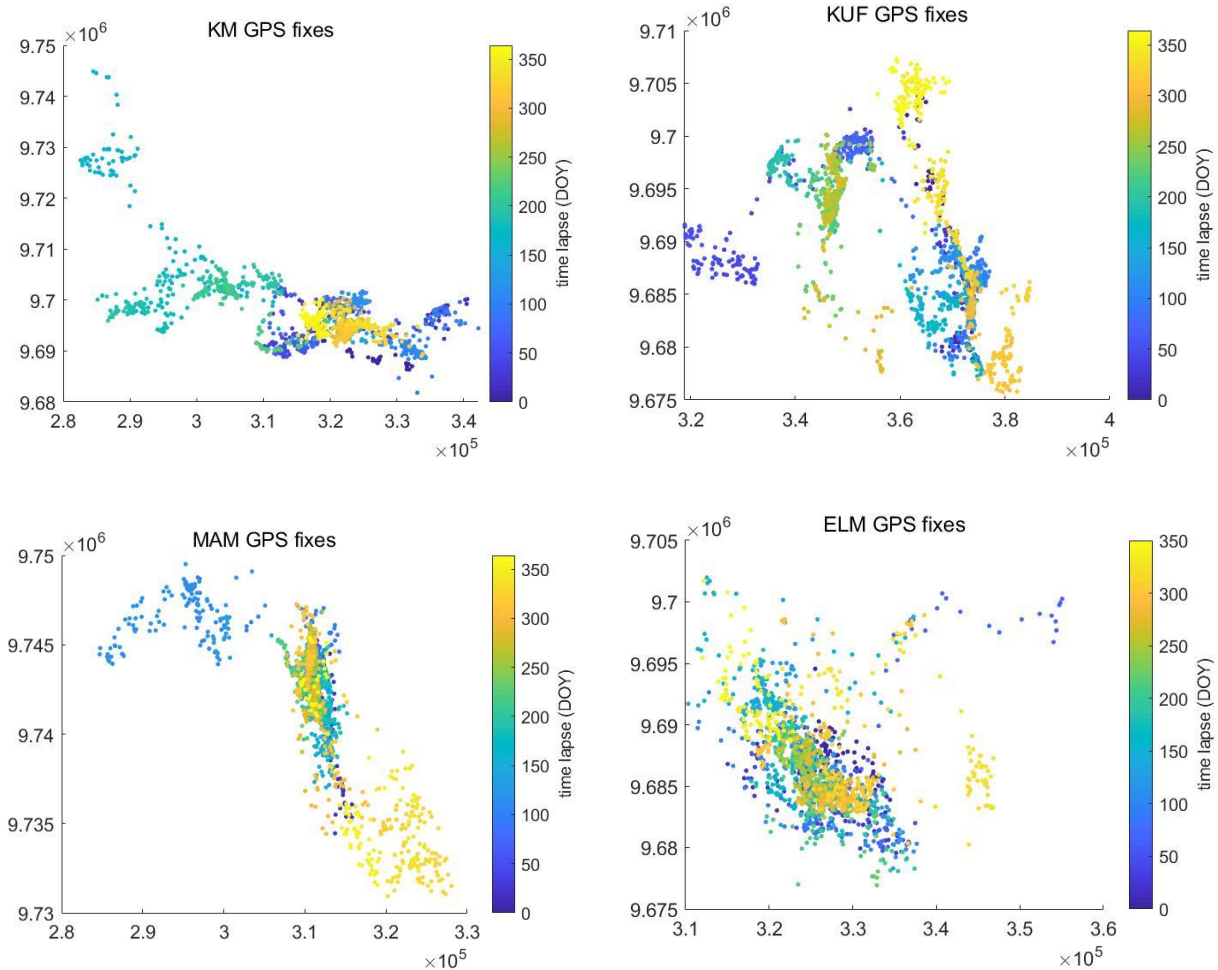


Figure 6 GPS fixes of four collared elephants in 2015 displayed in UTM Zone 37S coordinate system. KM, KUF, MAM and ELM are names of elephants.

2.5. Ancillary data

Ancillary data include Planet Scope Imagery, Digital Elevation Model (DEM) with 30 m spatial resolution, marshes map of Kenya in 2010, and surface water maps in the Amboseli ecosystem provided by the International Livestock Research Institute (ILRI) and Joint Research Centre (JRC) of European Commission acquired in 2012 and 2015, respectively. Detailed information about these datasets is described as follows,

Planet Scope ortho Scene Product consists of four bands, blue (455 nm-515 nm), green (500 nm-590 nm), red (590 nm-670 nm), and near-infrared bands (780 nm-860 nm), with a high spatial resolution of 3.125 m (Michael et al., 2018). All the Planet Scope ortho tile images acquired for the periods from July 17, 2016 to July 23, 2016, and from December 15, 2016 to December 20, 2016 over the Amboseli ecosystem were downloaded from the Planet Home (<https://www.planet.com/>) and used for assisting the visual interpretation of training and test dataset. Figure 7 shows the Planet Scope image (a small part of the Amboseli ecosystem) acquired on July 24, 2016.

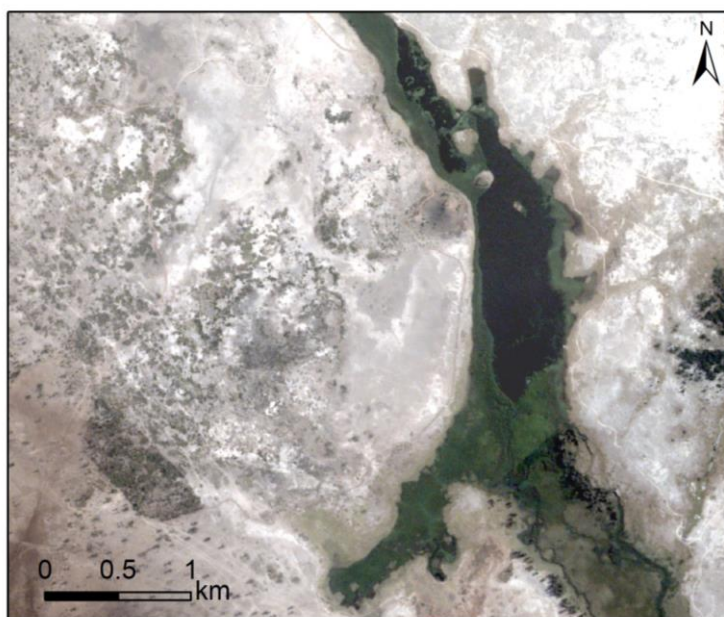


Figure 7 True color composite of Planet scope image (a subset of the study area) acquired on July 24, 2016

Digital Elevation Model (DEM) with 30 m spatial resolution of Kenya and Tanzania in 2015 were downloaded from the Regional Centre for Mapping of Resources for Development Geoportal (RCMRD) (<http://opendata.rcmrd.org/datasets>). The source of DEM is Shuttle Radar Topography Mission (SRTM), a project conducted by National Aeronautics and Space Administration (NASA) and the National Imagery and Mapping Agency (NIMA) of the U.S. Department of Defense (Farr & Kobrick, 2000). The vertical accuracy of the DEM data is 15 m. Surface water is closely linked with the terrain as water usually flows from higher places to the surrounding lower places (Huang et al., 2018). Therefore, integrating DEM data into the image classification may provide supplemental information for mapping the surface water.

Marshes map of Kenya in 2010 was downloaded from World Resources Institute (<https://www.wri.org/resources/data-sets>). The marshes map of Kenya is vector data in the form of polygon feature class with a geographic coordinate system of GCS WGS 1984. Surface water maps (include rivers, marshes) were provided by International Livestock Research Institute (ILRI) and JRC surface water

products were derived by (Pekel, Cottam, Gorelick, & Belward, 2016) using Landsat satellite images. Figure 8 shows the surface water maps provided by the ILRI. The rivers and marshes maps of Kenya provided by ILRI are vector data in the form of line and polygon feature class, respectively. Figure 9 shows the surface water maps provided by JRC. The surface water product downloaded from the JRC is in TIFF format with a spatial reference of GCS WGS 1984. The surface water product provided by JRC is extracted from Landsat satellite images with a spatial resolution of 30 m.

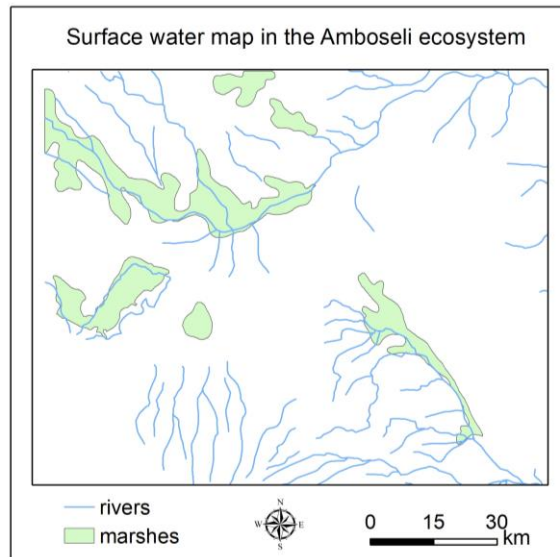


Figure 8 Surface water map in the Amboseli ecosystem provided by ILRI.

Surface water maps in the Amboseli ecosystem

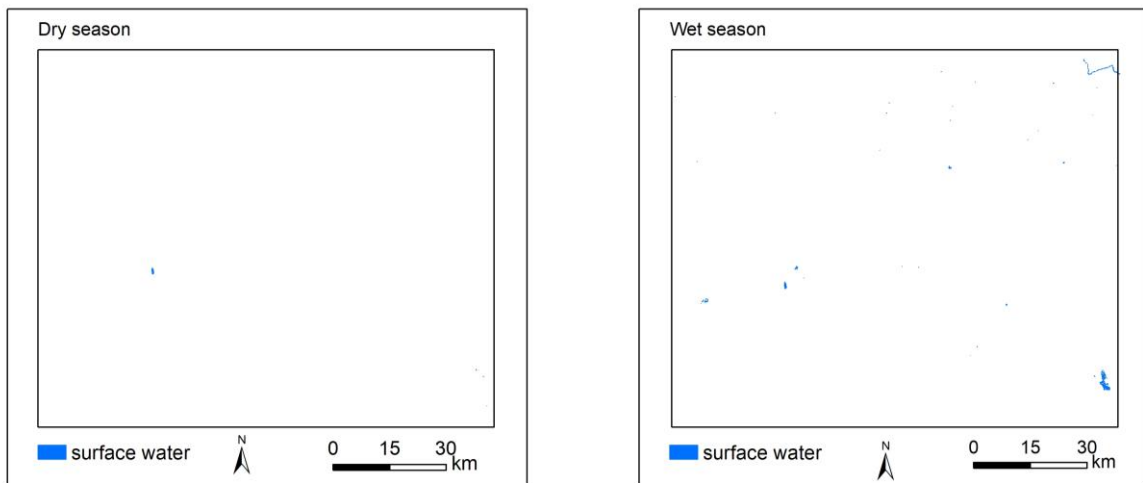


Figure 9 Surface water maps in dry season (left) and wet season (right) over the Amboseli ecosystem provided by JRC.

3. METHODOLOGY

This chapter mainly explains the surface water mapping using random forest algorithm and the method to analyze the relationship between the dynamic distribution of surface water and the elephant movements. This chapter consists of five sections. Section 3.1 describes the preprocessing procedures, indices calculation of Sentinel-1, Sentinel-2 and the processing of ancillary data. Section 3.2 mainly illustrates the processes of mapping the surface water using the random forest algorithm. This section elaborates the details of training and test data selection, input variables, parameter tuning, accuracy assessment and statistical tests of random forest methods. Section 3.3 describes the way to compare surface water products. Section 3.4 illustrates the steps to quantify and compare the distance of elephants to the nearest surface water. Section 3.5 demonstrates the method to analyze the movement speeds of elephants.

3.1. Data processing

3.1.1. Sentinel-1 radar data processing

Sentinel-1 IW mode Level-1 GRD products were selected for this study since dual-polarization was observed to perform better than single-polarization (Irwin et al., 2018). GRD products acquired on July 6, 2016, December 9, 2016 and December 14, 2016 over the Amboseli Ecosystem were downloaded from the European Space Agency Open Access Hub (<https://scihub.copernicus.eu/>). Images acquired in July is a representative of the dry season, while image obtained in December represents wet season because the precipitation is extremely low in July and relatively high in December respectively. Satellite images acquired on these dates are selected for this study to make the acquiring dates of images from different sensors, Sentinel-1 and Sentinel-2 as close as possible. After downloading and subsetting the radar imagery with the boundary of the study area, several procedures were conducted in SNAP 6.0 software to process the data, these procedures include,

- 1) Calibration. Calibration aims to transform the simple digital numbers into absolute normalized cross section (NRCS), also known as sigma naught (σ^0). Thus the pixel values can be linked with the backscatter. The type of calibration selected in this study is Sigma0.
- 2) Linear to/from dB. dB is the ratio of two values on a logarithmic scale. Converting the images in dB formats can get the radar backscatter and makes the histogram easier to work with and more comfortable to visualize because the difference between high and low backscatter is relatively larger in logarithm scale compared with a linear scale. The conversion of dB was performed from

$$\sigma_{dB}^0 = 10 \cdot \log_{10} \sigma^0 \quad (3.1)$$

where σ_{dB}^0 expresses the backscatter in dB, σ^0 expresses the sigma naught (Omar, Misman, & Kassim, 2017).

- 3) Convert band. Right click the virtual bands can convert “on the fly” bands into actual bands that could be saved or stored in the file.
- 4) Speckle filter. The speckle can decrease the quality of images and thus makes it more difficult for image interpretation. Refined Lee filter was adopted to reduce the speckle filter since it has been suggested that this filter can reduce the noise and maintain the boundary of open water as well (Morandeira, Grimson, & Kandus, 2016).
- 5) Terrain correction. Ranger doppler terrain correction in software SNAP 6.0 was used to compensate the geometric distortions caused by foreshortening and layover effects that resulted

from relief displacement. A Shuttle Radar Topography Mission (SRTM) 3Sec digital elevation model will be automatically downloaded from the Internet to remedy these distortions, and the bilinear interpolation was selected as DEM Resampling Method. UTM Zone 37 S was specified as the map projection.

- 6) Mosaicking. Mosaicking was carried out to combine two or more images into a larger one as the study area covered more than one image. Several parameters needed to specify while mosaic images in SNAP 6.0. Default projection, UTM/WGS84 was selected as the coordinate reference system. Also, the mosaic boundary was manually adjusted according to the boundary of the Amboseli ecosystem. Pixel size X and Pixel size Y both were set as 10 m.
- 7) Feature extraction. This step consists of indices calculation and grey-level co-occurring matrix (GLCM) because the indices and radar-derived texture have been observed to provide valuable information for land cover image classification in (Abdikan, Sanli, Ustuner, & Calò, 2016; Clerici, Valbuena Calderón, & Posada, 2017; Vreugdenhil et al., 2018). Band math tool in SNAP 6.0 was used to calculate the polarized ratio and average of VV and VH according to the formulas described in Table 2. GLCM was calculated from the texture analysis tool, and 5*5 was selected as the window size, other parameters remained as default values.

Table 2 Formulas of indices in Sentinel-1 radar data.

Indices	Formula	Reference
Polarized ratio (VH to VV)	VH/VV	(Vreugdenhil et al., 2018)
Average of VV and VH	$(VV+VH)/2$	(Abdikan et al., 2016)

All the processed bands were exported as GeoTIFF format for further analysis in ArcMap 10.6.1 and RStudio.

3.1.2. Sentinel-2 optical data processing

Sentinel-2 level-1C satellite images obtained on 23 July 2016 and 20 December 2016 over the Amboseli Ecosystem were downloaded from the European Space Agency Open Access Hub (<https://scihub.copernicus.eu/>). Four images on each date are needed to cover the whole study area. After downloading the images, cloud masks were generated from the Fmask 4.0 software developed by (Qiu, He, Zhu, Liao, & Quan, 2017; Frantz et al., 2018). The generated cloud masks are demonstrated in Figure 10 and Figure 11. Parameters include cloud probability threshold, potential false positive cloud and result dilations were specified as default values in Fmask 4.0. The atmospheric correction had been conducted on Level-1C satellite images in Sen2Cor, after which, all the image bands with JPEG2000 format were exported as TIFF format in ArcMap 10.6.1. Then they were subset based on the boundary. Next, pixel-based mosaicking was implemented in ENVI 5.5. After pre-processing of Sentinel-2 satellite images, indices were calculated according to the formulas in Table 3 in the band math tool of ENVI 5.5 software. Then, all the image bands were exported as TIFF format file for further analysis in RStudio.

Table 3 Formulas of water and vegetation indices in Sentinel-2 satellite images (Green = Band 3, Red = band 4, near infrared (NIR)= Band 8).

Indices	Abbreviations	Formulas	Reference
Normalized Difference Water Index	NDWI	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	(McFeeters, 1996)
Normalized Difference Vegetation Index	NDVI	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	(Rouse, Hass, Schell, & Deering, 1972)

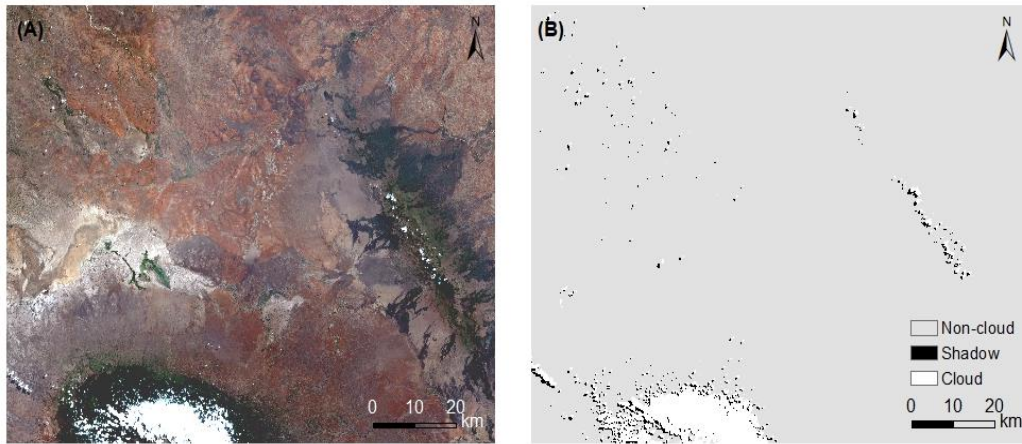


Figure 10 True color composite (R, G, B =Band4, 3, 2) of Sentinel-2 image acquired on July 23, 2016 (A), and the corresponding cloud and shadow mask generated from Fmask 4.0 software (B).

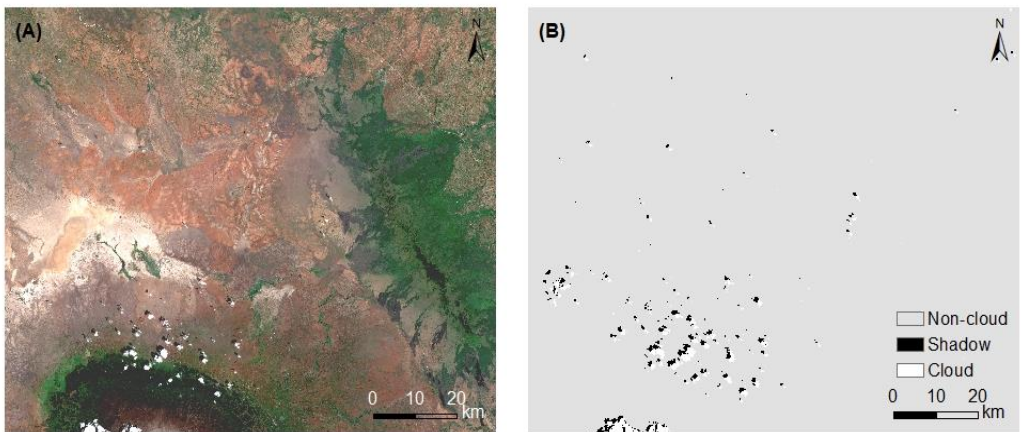


Figure 11 True color composite (R, G, B =Band4, 3, 2) of Sentinel-2 image acquired on December 20, 2016 (A) and the corresponding cloud and shadow mask generated from Fmask 4.0 software (B).

3.1.3. Ancillary data processing

Ancillary data include Planet Scope Imagery, DEM, marshes map of Kenya in 2010, and surface water maps provided by ILRI and JRC acquired in 2012 and 2015, respectively. The description and processing of these datasets are as follows,

Planet Scope ortho tile images were displayed in software ArcMap 10.6.1 for assisting the selection and interpretation of training and test points. DEM in 2015 was subsetting, resampled to 10 m and re-projected into UTM Zone 37 S coordinate system in ENVI 5.5. Slope and aspect were generated from DEM in ArcMap 10.6.1.

Marshes map of Kenya in 2010 was projected to the UTM Zone 37 S coordinate system then subset according to the boundary of the study area, after which boundaries of the marshes were adjusted manually based on the Sentinel-2 images in ArcMap 10.6.1. The marshes map was integrated into the surface water maps derived from Sentinel images. Surface water maps (include rivers and marshes) provided by ILRI was re-projected to UTM Zone 37 S coordinate system, subsetting to the boundary of the study area. JRC surface water products with TIFF format were converted into polygon feature class using the same software.

3.2. Mapping surface water using random forest

Random forest is a machine learning method that consists of a large number of decision trees (Breiman, 2001). Random forest trains the classifier to predict the class (label) of each object (pixel or polygon). The class of an object is labeled as the class that have received the most votes from all trees in the forest. Trees are established from a bootstrap sample of training dataset, and the splitting at the node of each tree is based upon a random selection of a subset of variables (Balzter, Cole, Thiel, & Schullius, 2015). Random forest consists of a certain number of trees (Ntree), and every tree contains several nodes. Each tree represents a classifier and votes for the labels of the class. The final label of a particular class type is determined by the class that had received the most votes from the trees.

Random forest is a robust machine learning algorithm that has been widely and being increasingly used in image classification and regression (Horning, 2010). Random forest algorithm has many advantages over other machine learning algorithms, like support vector machines (SVM) and neural network. First of all, the random forest is not very sensitive to parameters, and it is not difficult to decide which parameter to choose (Breiman, 2001). Also, the random forest can yield information about the relative importance of each input variable (variable importance), which makes this approach more efficient and accurate (Horning, 2010). Moreover, a random forest is not sensitive to missing data and it can predict missing data effectively as well as maintain accuracy even if a large part of the data are missing (Breiman, 2001; Kulkarni, 2013). Finally, the random forest is not computationally intensive classifier (Gislason, Benediktsson, & Sveinsson, 2006). Thus, the random forest algorithm is selected for surface water mapping in this study.

Random forest classification was implemented in RStudio through running packages, “sp”, “rgdal”, “raster”, “randomForest”, “ggplot2”, using different datasets and their derived variables as the inputs of models. Six models were established to derive the surface water maps.

3.2.1. Training and test datasets collection

Ideally, a stratified random sampling strategy is suggested to reduce bias for selecting training and test points. However, random sampling is not feasible since the High-Resolution Image (Planet Scope) cannot cover the whole study area. Therefore, purposive sampling was adopted in this study. 160 points were selected and interpreted as training and test datasets for the classification of Sentinel images acquired in July and December, respectively. Among the 160 points, 80 points were labelled as water, and the rest are assigned as Non-water. True color composite image was generated through assigning Band 4, Band 3 and Band 2 to the R, G, B colors for display. In true (natural) color composite, objects on the satellite images closely correspond with what human being observed in the real world. Hence, water usually appears blue and black on satellite images with true color composite. Sometimes, the water looks brown or green depends on the proportion of suspended sediments or phytoplankton. NDWI calculated from Sentinel-2 images were classified with a relatively strict threshold value, 0.3 ($NDWI > 0.3$ represents water), proposed by (McFeeters, 2013), for assisting in finding the location of water since the proportion of water is relatively smaller compared with Non-water class and it increases the difficulty in locating the surface water. Non-water covers different land covers, including forest, grassland, cropland, built-up and bare land, to make the training and test points as representative as possible. Training and test datasets were converted to KML format file and the labels of all the points were double-checked in Google Earth Pro. Selected training and test points for random forest image classification are shown in Figure 12.

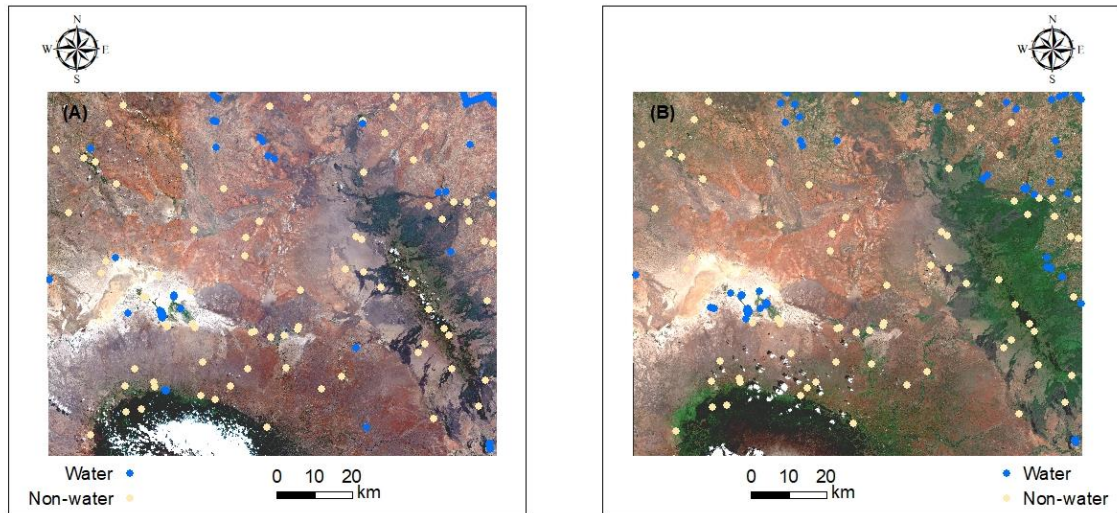


Figure 12 Distribution of training and test points in the Amboseli ecosystem during the dry season (A) and wet season (B).

3.2.2. Input variables of random forest

Random forest algorithm measures the importance of input variables in the prediction of the response variable (Janitzka, Tutz, & Boulesteix, 2014). Therefore, the variable importance can be used to select suitable variables or features with higher predictability. Mean Decrease Gini (MDG) is a coefficient that measures the variable importance through calculating the average decrease in node impurity (increase in node purity) of splits at nodes using a given variable (Han, Guo, & Yu, 2017). Gini coefficient is a value that varies from 0 (perfect equity) to 1 (perfect inequity) that it was widely been used to measure the distribution of income or wealth. If the Gini coefficient is small, the distribution is homogeneous. The more the Gini coefficient decrease, the larger the contribution of the variable in the homogeneity of the nodes. Hence, the higher the MDG value, the greater the importance of the variable. The importance score of MDG is described as,

$$VI(x^j) = \frac{1}{n} [1 - \sum_{k=1}^n Gini(j)^k] \quad (3.2)$$

where n expresses the number of tree nodes based on feature x^j , j is the ranking in descending order of the features, $Gini(j)^k$ is the k^{th} Gini index of feature x^j . Eq.(3.2) is adapted from (Wang, Yang, & Luo, 2016).

The input variables of each model can be found in Table 4. Sentinel-1 radar data consists of five variables, VH, VV, VH or VV, Average and VH correlation. Sentinel-2 optical data consists of six variables, including Band2, Band3, Band4, Band8, NDWI and NDVI. DEM data include elevation, slope and aspect.

Table 4 Input variables of random forest models using different datasets.

Model	Input variables					Number of variables
	Sentinel-1 (July)	Sentinel-2 (July)	Sentinel-1 (December)	Sentinel-2 (December)	DEM	
A	√				√	8
B		√			√	9
C	√	√			√	14
D			√		√	8
E				√	√	9
F			√	√	√	14

Variable importance can be used for feature selection for the random forest models since it measures the importance of each input variable. The variable importance of each variable in each random forest model can be referred to in Figure 13. In Figure.13(A), the most important variable is the average of VV and VH. While for model D, the most important variable is VH. For random forest model using Sentinel-2 optical data, NDWI is the most important variable in model B and model E. Similarly, for classification models using the combination of Sentinel-1 radar and Sentinel-2 optical data, NDWI still contributes the most. Due to the advantage of measuring the variable importance of random forest, variables with low variable importance, such as slope and aspect could be moved out of the models. However, the accuracy and kappa coefficients of classification using Sentinel-1 radar data and Sentinel-2 optical are compared in this study and the importance of these two variables varies, thus these two variables were not deleted.

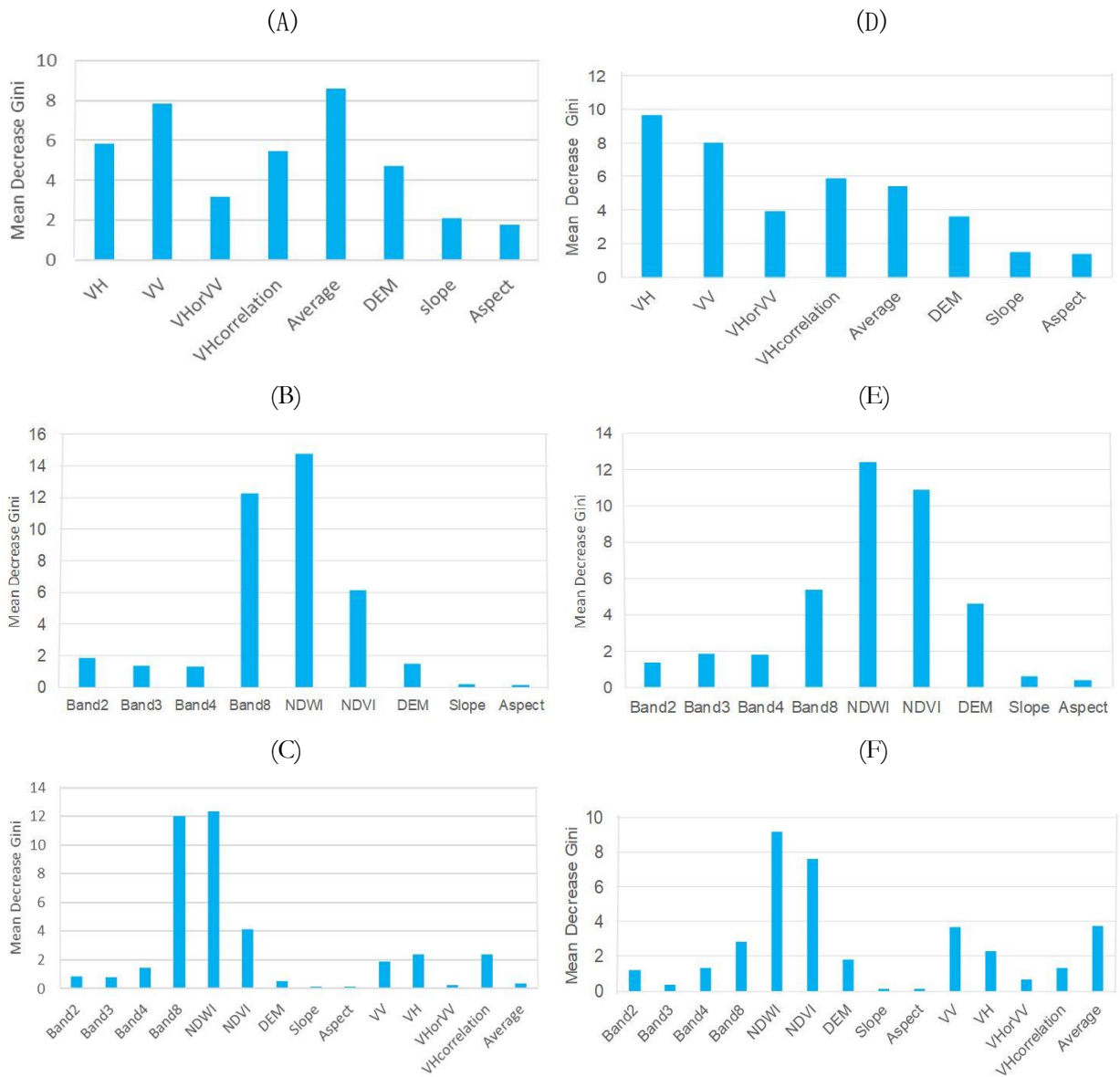


Figure 13 Variable importance of random forest classification results using different datasets.

(A)-(C) represent the variable importance of image classification in July 2016 (dry season) using Sentinel-1 data (A), Sentinel-2 data (B), and the combination of both (C). (D)-(F) represent the variable importance of image classification using different data in December 2016 (wet season) using Sentinel-1 data (D), Sentinel-2 data (E), and the combination of both (F).

3.2.3. Parameter tuning of random forest

Parameter tuning is a critical part for improving the performance of random forest models through searching for the “optimal” values for parameters. Two important parameters, “mtry” and “ntree”, were tuned in this study. “mtry” is the number of randomly selected variables when splitting at tree node. The “mtry” is defaulted as the square root of the number of input variables in RStudio. A grid search was adopted for the selection of “mtry”. Usually, points in the grids represent the combination of two parameters. The grid search here is a linear search for trying all the possible values of “mtry” in this study. Grid search with repeated 10-fold cross-validation with ten repeats was implemented for the parameter estimation of “mtry”. Figure 14 shows the effect of “mtry” on accuracy. The values of “mtry” parameter range from 1 to the number of input variables. Parameter “mtry” can be determined when the cross-validation accuracy reaches the maximum. Table 5 demonstrates the selected values of “mtry” of different random forest models. “ntree” represents the number of trees established in the forest. Those unselected samples were considered out of bag (OOB) observations (Janitza & Hornung, 2018). The OOB error is defined as,

$$e = \frac{1}{N} \sum_{i=1}^N \frac{1}{|c^{-i}|} \sum_{b \in c^{-i}} L(Y_i, \hat{f}^b(x_i)) \quad (3.3)$$

Where N expresses the sample size, c^{-i} is the indices that not belong to the set of bootstrap sampling b , $|c^{-i}|$ is the number of samples that not belongs to the set b , L is the loss function, Y_i is the true outcome, $\hat{f}^b(x_i)$ is the bias, adapted from (Goldstein, Polley, & Briggs, 2011).

The OOB observations can be used to evaluate the performance of the random forest model, resulted in the so-called OOB error rate. Lower OOB error rate means higher OOB accuracy and higher classification accuracy in random forest models based on the training sets. Previous research has claimed that OOB error can be used as an unbiased estimator for assessing the accuracy of random forest models (Goldstein et al., 2011). However, using the OOB error to estimate the final prediction error of the random forest model can be biased in image classification (Janitza & Hornung, 2018). Thus, the kappa coefficient instead of OOB error rate was selected to assess the overall accuracy of image classification using random forest algorithm. OOB error stabilized at a certain number of trees. 500 trees (default value) were first constructed in each random forest model. Figure 15 illustrates the number of trees and the corresponding OOB error rate. OOB error rate fluctuates when the number of trees is relatively small. While the number of trees is sufficiently large enough, the OOB error converges. Thus, the number of trees can be selected after the stabilization of OOB error. In Figure 15(A), the OOB error stabilizes when the number of trees grows at about 100. Hence, values that are larger than 100 can be decided as the number of trees of that model. Parameter “ntree” picked in each random forest model can be referred to Table 5.

Table 5 Tuned parameters for random forest models used to map surface water in the Amboseli ecosystem.

Models	Tuned parameters	
	Mtry (number of randomly selected variables)	ntree (number of trees)
A	2	150
B	2	50
C	5	250
D	1	250
E	3	350
F	4	200

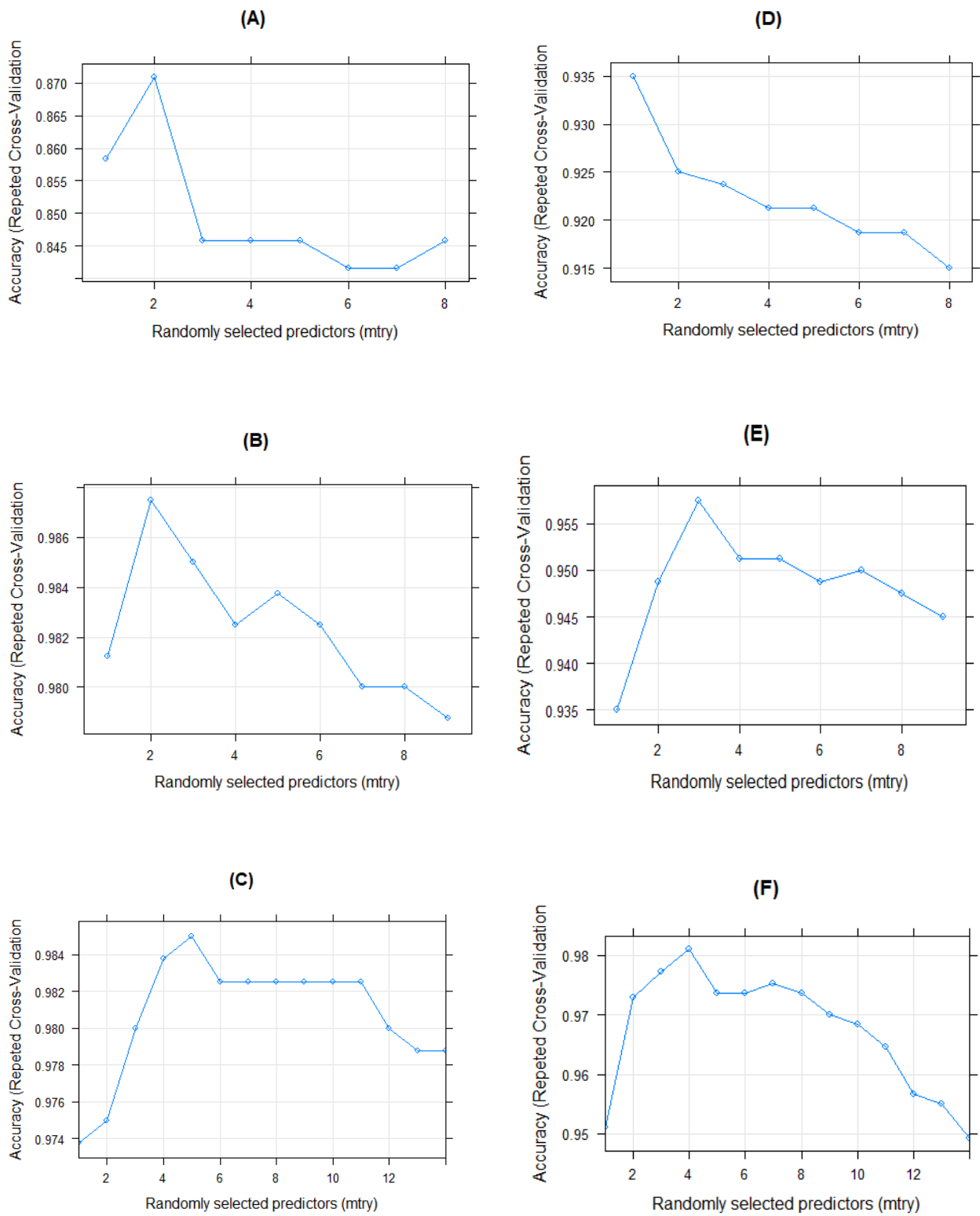


Figure 14 Tuning of “mtry” parameter of six different random forest models using different imagery.

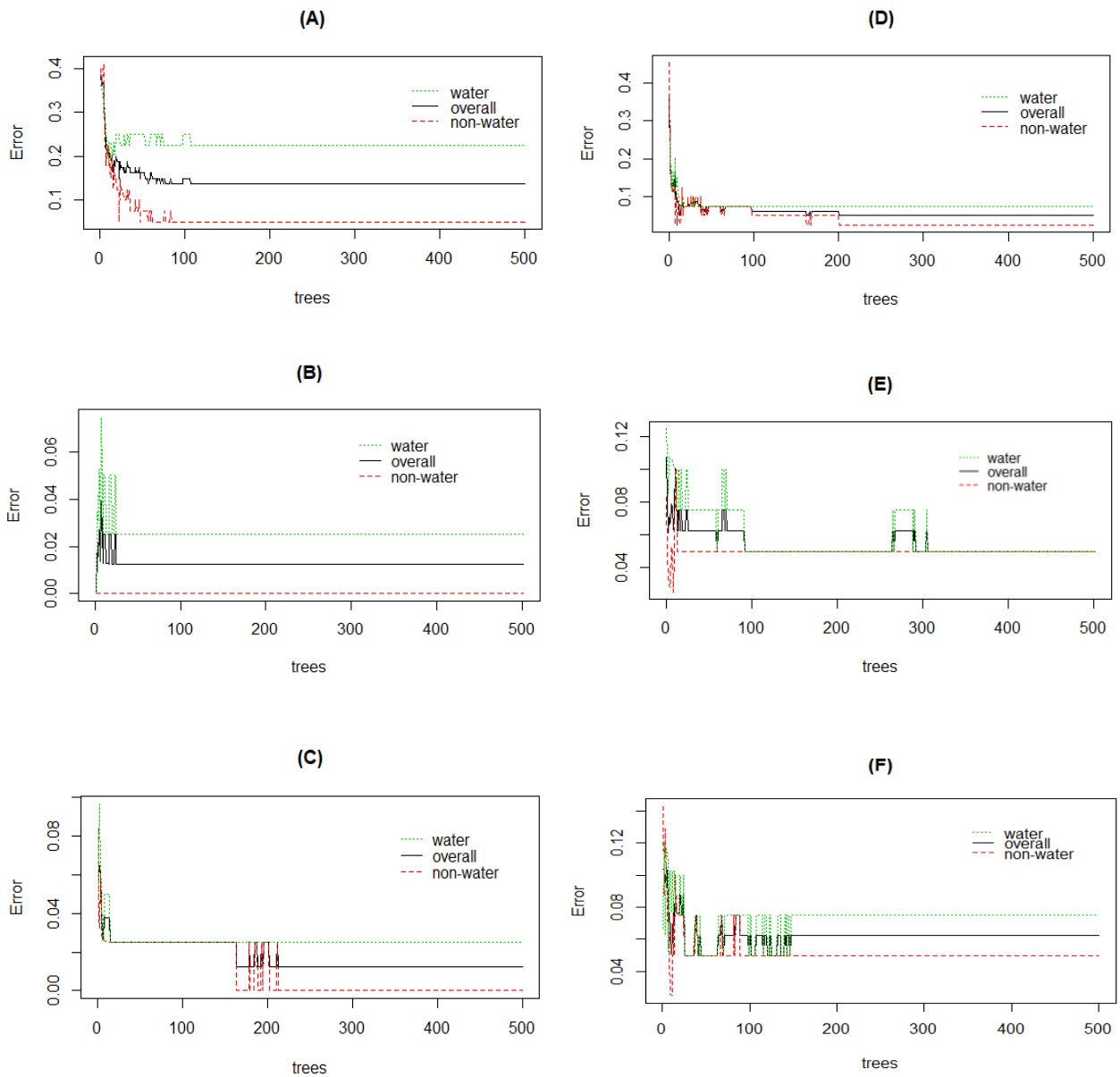


Figure 15 Effects of number of trees on the OOB error rate based on the training data of six random forest models. Black line represents the overall OOB error rate of all the classes in random forest classification. While green and red lines represent the OOB error rate of water and Non-water class in the random forest models. The horizontal axis represents the parameter “ntree”, number of trees while the y axis represents the OOB error rate.

3.2.4. Accuracy assessment

The accuracy of surface water maps was assessed by the confusion matrix and the kappa coefficients. The confusion matrix is a simple matrix of the predicted class against true class in reality (reference data collected from fieldwork) at specific locations (Foody, 2002). This matrix describes the classification accuracy and errors committed or omitted in classification maps, which helps identify which class was wrongly classified to other classes. Confusion matrix can evaluate the performance of a classification algorithm through the producer's accuracy, user's accuracy and overall accuracy. Producer's accuracy describes the accuracy from the perspective of the producer (map maker). The map users focus on thinking how good the ground truth can be represented by maps. Thus, correctly classified reference points (or pixels) in each class divide by total reference points in that particular class is the producer's accuracy. On the contrary, user's accuracy demonstrates the accuracy from the perspective of map user. The users always think how reliable the maps are and they focus on understanding how good the map can represent the reality. Therefore, correctly classified reference points in one class divide the row total (total reference points labeled as that class in maps). Overall accuracy illustrates the portions of the correctly classified reference points. Correctly classified points for all the classes (elements in the diagonal of the matrix) divide the total reference points is the overall accuracy. Although confusion matrix has been widely used for accuracy assessment in image classification, one major problem, namely chance agreement, is inevitable. Chance agreement problem refers that for some features that assigned to the correct class purely by chance (Foody, 2002). Cohen's kappa coefficient was proposed to remedy limitations caused by chance agreement in overall accuracy. Cohen's kappa coefficient measures the proportion of agreement after excluding chance agreement (Cohen, 1960). Cohen's kappa coefficient describes how good a classifier can predict the true class label compared with the randomly assigned class label. The value of the kappa coefficient varies from -1 to 1. The classification performance can be evaluated as poor ($\text{kappa} < 0$), slight ($\text{kappa}: 0.01-0.20$), fair ($\text{kappa}: 0.21-0.40$), moderate ($\text{kappa}: 0.41-0.60$), substantial ($\text{kappa}: 0.61-0.80$) and almost perfect ($\text{kappa}: 0.81-0.99$) (Viera & Garrett, 2005). Test datasets (reference points) for accuracy assessment was interpreted from Planet Scope satellite images. 40 reference points of each class, water and Non-water, were interpreted for the accuracy assessment of the classified thematic maps.

3.2.5. Accuracy comparison

Each random forest model was iterated for five hundred times. Five hundred kappa coefficients were generated from each model. Kappa coefficients were claimed to be normal distribution when the sample size is large based on the central limit theorem (Vanbelle, 2017). Two-sample t-test ($\alpha = 0.05$) was adopted to test if any significant difference is existing in the mean kappa coefficients of different random forest models because two-sample t-test is considered one of the most widely used statistical tests that compare if the average difference between two groups is statistically significant.

3.3. Comparing surface water products

3.3.1. Comparing the number and area of surface water

After mapping the surface water in the Amboseli ecosystem using Sentinel data during the dry and wet season of 2016. Two maps with highest kappa coefficients in dry and wet season are selected for further analysis. The TIFF format (Sentinel and JRC water products) files were converted into polygon feature class in ArcMap 10.6.1. The number and area of water bodies can be inquired in the attribute table of the feature class. For calculating the number of water bodies, water bodies that locate within 10 m from each other were considered as the same water body. Rivers and lakes (or small water bodies) were derived from the satellite images while marshes was integrated into the Sentinel surface water maps from external maps layer.

3.3.2. Comparing the mean distances to surface water

Numerous methods can be used to compare raster datasets. Similarity and dissimilarity measures have been widely used to compare the maps in previous studies (Lavigne, Ricci, Franck, & Senoussi, 2010; Ray & Burgman, 2006). Similarity measures including but not limited to Pearson correlation coefficient, Cosine similarity and Euclidean distance (Jyoti et al., 2015). Histogram comparison between paired pixels also has been investigated before (Wells et al., 2003). In addition, the statistical test was investigated to compare the mean difference between different pixel values (Lavigne et al., 2010). This study combines the distance based and statistical test approach to compare the different surface water products. A pixel-by-pixel comparison of the surface water maps is impractical in this study since the study area covers a large area of approximately 11000 km² and consists of a huge number of pixels. Euclidean distance was calculated for the three surface water products, which include Sentinel, JRC and ILRI. 3000 random points were generated using the “Create random points” tool in ArcMap 10.6.1. After the creation of these random points, the tool “Extract Multi Values to points” was adopted to extract the distance to the surface water. Then, one -way ANOVA was conducted in RStudio to test the difference in the mean distances of random points to surface water during the dry and wet season. The assumption was made about the normal distribution of the distances extracted from different surface water products since the 3000 points are randomly created. Even if the distances are not normally distributed, the previous study has observed that the ANOVA is still valid even if the data is not normally distributed (Blanca, Alarcón, Arnau, Bono, & Bendayan, 2017). After conducting ANOVA ($p < 0.05$), not all mean distances are equal. Tukey’s Honestly Significant Difference test, a post hoc test was implemented to identify the distances groups means that have a difference. The significance level for the test is 0.05.

3.4. Quantifying and comparing the distance of elephants to the nearest surface water

The distance of elephant to the nearest surface water based on different surface water products (i.e. Sentinel, JRC and ILRI) were extracted to the GPS fixes. The distance of elephants to the nearest surface water in the Amboseli ecosystem during different seasons using different surface water products were implemented with the Shapiro Wilk test ($\alpha=0.05$). Results showed that the distances of elephants are not normally distributed ($p < 0.05$). Thus Mann-Whitney U test ($\alpha=0.05$) was adopted to test the difference in the distance to the nearest surface water between the dry and wet season. The difference in the distances of elephants to the nearest surface water extracted from the three surface water products was tested separately.

3.5. Analyzing the relationship between elephant movement and the surface water availability

3.5.1. Calculating the movement speeds of elephants

GPS fixes were re-projected to UTM Zone 37S projected coordinate system and fixes with missing values were deleted before any further analysis. GPS fixed obtained in 2015 were chosen for this study. GPS fixes obtained in July and August were classified as dry season, whereas those acquired in November and December were categorized as wet season. A total of 2564 GPS fixes was used in this study (as shown in Table 5). The percentages of missing GPS fixes of elephant KM, KUF, MAM, and ELM are 8.1%, 9.2%, 16.5% and 6.5%.

The movement speed of elephants was calculated using “PrepData” function in the package “moveHMM” in Rstudio. Two procedures were conducted to calculate the movement speed of the elephants. First of all, the step length, the Euclidean distance between two successive GPS fixes was calculated. Then the step length was divided by time intervals between two successive fixes. The unit of the calculated speed is in kilometers per hour (km/h). The calculated speeds were checked for potential outliers using a threshold

7km/h because the maximum movement speed of the African elephant is about 7km/h. Table 6 demonstrates the basic statistics of calculated speed during the dry season and wet season.

Table 6 Descriptive statistics of elephant movement speed (Unit: km/h) during the dry and wet seasons.

Season	Minimum	Maximum	Mean	First Quartile	Median	Third Quartile
Dry	0.002	2.21	0.34	0.11	0.23	0.46
Wet	0.002	1.49	0.27	0.10	0.20	0.35

3.5.2. Analyzing the relationship between the speed of the elephant and the surface water availability

After the calculation of the movement speed of each elephant. A scatter plot of two variables, distance to the nearest surface water and speed of elephant movement, was not display any particular pattern. Therefore, the distance to water was classified into three categories: 1) close to water (< 3 km); Medium to water (3 km- 6 km); far from water (> 6 km), followed previous research conducted by Kummur et al (2011). Shapiro Wilk test ($\alpha=0.05$) was adopted to test the normality of the movement speed in each category. The null hypothesis of the Shapiro Wilk test is that the data was drawn from a normal distribution and the null hypothesis was rejected since all the p values in the test of each class are lower than 0.05. Then, a nonparametric test, Mann-Whitney U test ($\alpha=0.05$) was conducted to test the difference in the movement speeds of elephant among the three distance categories. Analysis of the relationship between the speeds of elephants and the surface water availability were implemented based on the Sentinel and ILRI surface water product. The JRC surface water product was not analyzed since the distance of elephants to the nearest surface water is too long.

4. RESULTS

The results mainly present the surface water maps in the Amboseli ecosystem and the relationship between the elephant movements and surface water availability. This chapter consists of four sections. Section 4.1 shows the surface water maps and confusion matrix. Section 4.2 illustrates the comparison of different surface water products. In section 4.3, the distance of elephants to the nearest surface water is demonstrated. The following section 4.4 illustrates the relationship between the speed of the elephant and the surface water availability.

4.1. Surface water maps

4.1.1. Surface water maps derived from Sentinel-1 radar data and DEM

Figure 16 shows the classification result in the Amboseli ecosystem during the dry season using Sentinel-1 data and DEM. The confusion matrix of the image classification was shown in Table 7. Among the 40 reference points for each class collected for accuracy assessment, 30 points of water class were correctly classified, and 31 points of Non-water class were correctly classified. The overall accuracy is 76.25% and the kappa coefficient is 0.53. The producer's accuracy of water and Non-water are 75% and 77.50%, respectively. The user's accuracy of water and Non-water are 76.92% and 75.61%, respectively. The producer's accuracy is the complement of the omission error, producers' accuracy plus the omission error equals to 100%; while the users' accuracy is the complement of the commission error, user's accuracy plus commission error equals to 100%. Therefore, the omission errors of water and Non-water are 25.00% and 22.50%, the commission errors of water and Non-water are 23.08% and 24.39%.

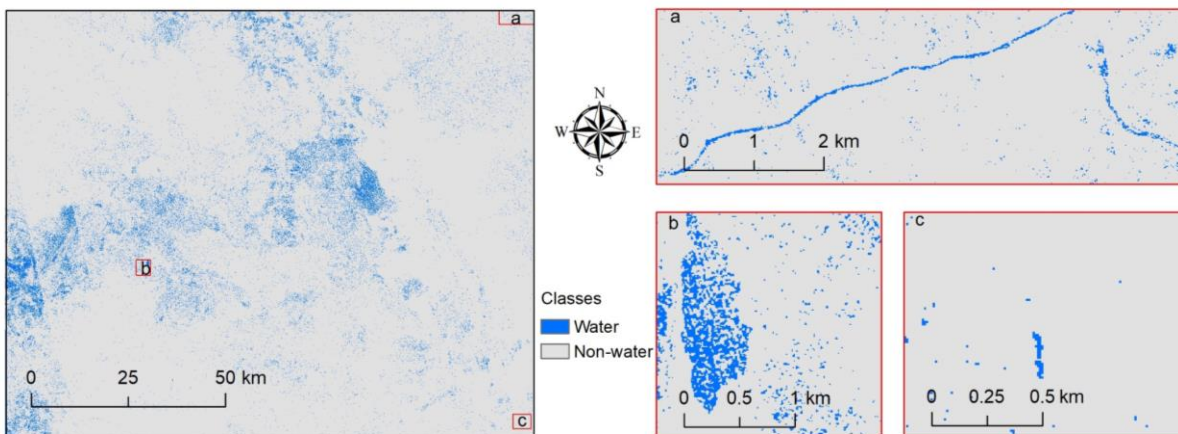


Figure 16 Surface water map of the Amboseli ecosystem during the dry season of 2016 based on Sentinel-1 radar imagery and DEM data.

Table 7 The confusion matrix of classification using Sentinel-1 radar data and DEM during the dry season.

Prediction	Reference		User's accuracy
	1	2	
1-Water	30	9	76.92%
2-Non-water	10	31	75.61%
Producer's accuracy	75.00%	77.50%	
	Overall accuracy		76.25%
	Kappa		0.53

Figure 17 shows the classification result in the Amboseli ecosystem during the wet season using Sentinel-1 radar data and DEM. The confusion matrix of the image classification was shown in Table 8. Among the 40 reference points for each class collected for accuracy assessment, 30 points of water class were correctly classified, and 33 points of Non-water class were correctly classified. The overall accuracy is 78.75% and the kappa coefficient is 0.57. The producer's accuracy of water and Non-water is 75.00% and 82.50%, respectively. The user's accuracy of water and Non-water are 81.08% and 76.74%, respectively. The omission errors of water and Non-water are 25.00% and 17.50%, the commission errors of water and Non-water class are 18.92% and 23.26%.

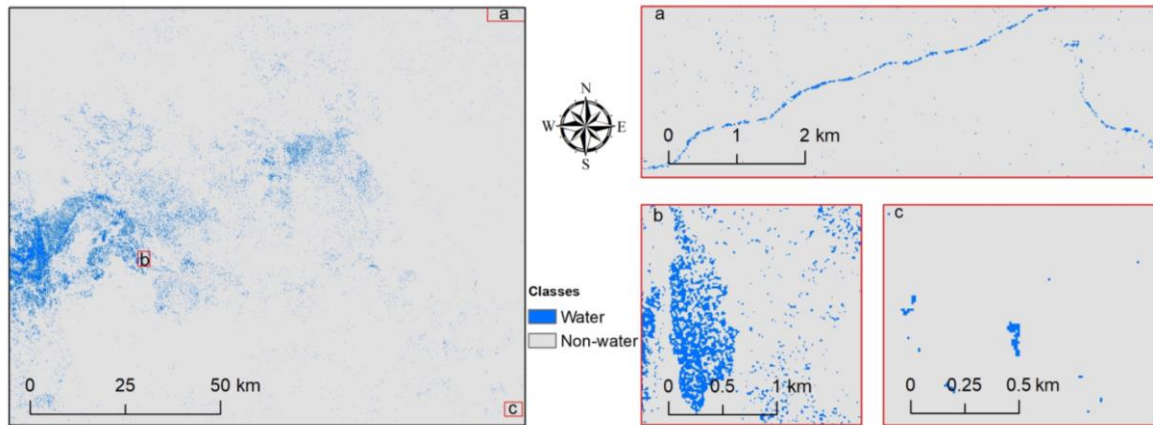


Figure 17 Surface water map of the Amboseli ecosystem during the wet season of 2016 based on Sentinel-1 radar imagery and DEM data.

Table 8 The confusion matrix of classification using Sentinel-1 radar and DEM data during the wet season.

Prediction	Reference		User's accuracy
	1	2	
1-Water	30	7	81.08%
2-Non-water	10	33	76.74%
Producer's accuracy	75.00%	82.50%	
	Overall accuracy		78.75%
	kappa		0.57

4.1.2. Surface water maps derived from Sentinel-2 optical data and DEM

Figure 18 shows the classification result in the Amboseli ecosystem during the dry season using Sentinel-2 optical data and DEM. The confusion matrix of the image classification was shown in Table 9. Among the 40 reference points for each class collected for accuracy assessment, 36 points of water class were correctly classified, and 40 points of Non-water class were correctly classified. The overall accuracy is 95.00% and the kappa coefficient is 0.90. The producer’s accuracy of water and Non-water are 90.00% and 100.00%, respectively. The user’s accuracy of water and Non-water are 100.00% and 90.91%, respectively. Therefore, the omission errors of water and Non-water are 10.00% and 0%, the commission errors of water and Non-water class are 0% and 9.09%.

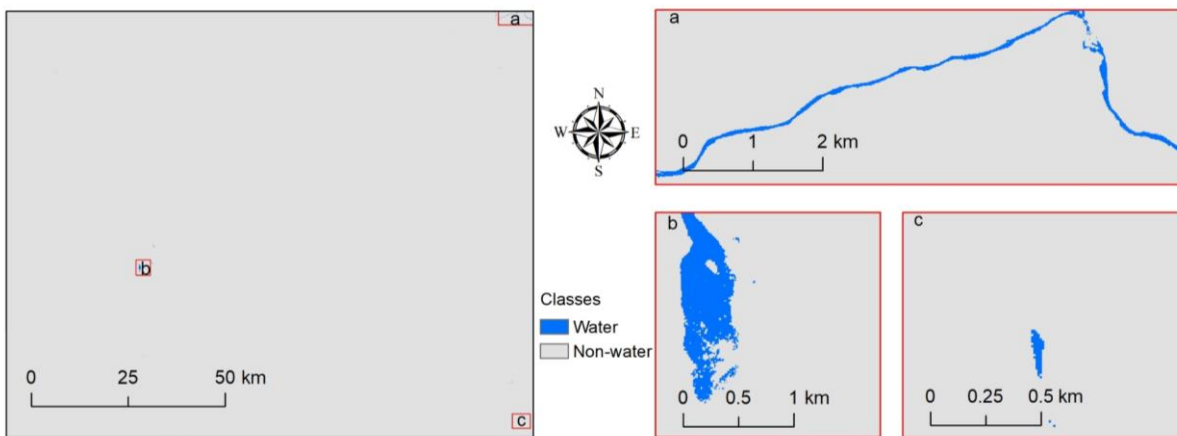


Figure 18 Surface water map of the Amboseli ecosystem during the dry season of 2016 based on Sentinel-2 optical imagery and DEM data.

Table 9 The confusion matrix of classification using Sentinel-2 optical data and DEM during the dry season.

Prediction	Reference		User's accuracy
	1	2	
1-Water	36	0	100.00%
2-Non-water	4	40	90.91%
Producer's accuracy	90.00%	100.00%	
	Overall accuracy		95.00%
	Kappa		0.90

Figure 19 shows the classification result in the Amboseli ecosystem during the wet season using Sentinel-2 optical data and DEM. The confusion matrix of the image classification is shown in Table 10. Among the 40 reference points for each class collected for accuracy assessment, 33 points of water class were correctly classified, and 40 points of Non-water class were correctly classified. The overall accuracy is 91.25% and the kappa coefficient is 0.83. The producer's accuracy of water and Non-water are 82.50% and 100.00%, respectively. The user's accuracy of water and Non-water are 100.00% and 85.11%, respectively. Therefore, the omission errors of water and Non-water are 17.50% and 0%, the commission errors of water and Non-water class are 0% and 14.89%.

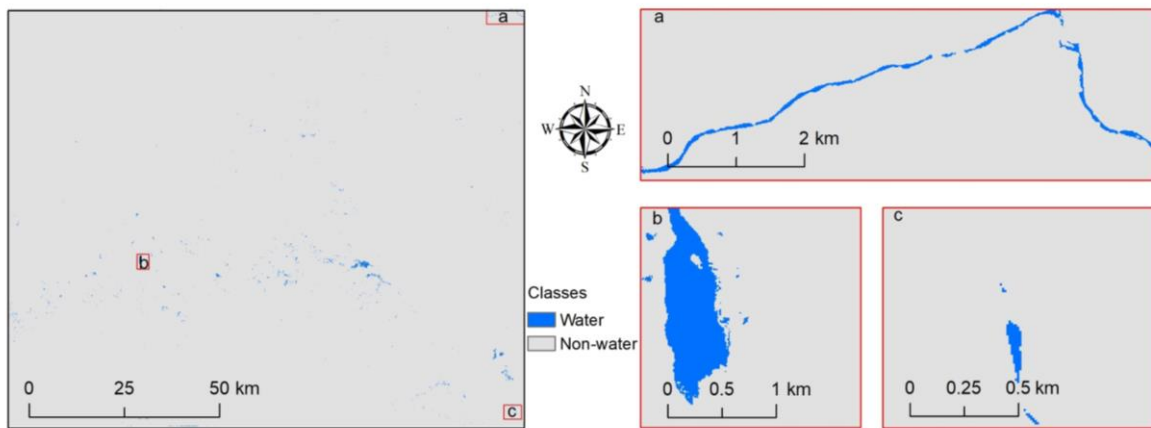


Figure 19 Surface water map of the Amboseli ecosystem during the wet season of 2016 based on Sentinel-2 optical imagery and DEM data.

Table 10 The confusion matrix of classification using Sentinel-2 optical and DEM data during the wet season.

Prediction	Reference		User's accuracy
	1	2	
1-Water	33	0	100.00%
2-Non-water	7	40	85.11%
Producer's accuracy	82.50%	100.00%	
	Overall accuracy		91.25%
	Kappa		0.83

4.1.3. Surface water maps derived from the combination of Sentinel-1, Sentinel-2 data and DEM

Figure 20 shows the classification result in the Amboseli ecosystem during the dry season using Sentinel-1, Sentinel-2 data and DEM. The confusion matrix of the image classification is shown in Table 11. Among the 40 reference points for each class collected for accuracy assessment, 36 points of water class were correctly classified, and 40 points of Non-water class were correctly classified. The overall accuracy is 95.00% and the kappa coefficient is 0.90. The producer’s accuracy of water and Non-water are 90.00% and 100.00%, respectively. The user’s accuracy of water and Non-water are 100.00% and 90.91%, respectively. The omission errors of water and Non-water are 10.00% and 0%, the commission errors of water and Non-water class are 0% and 9.09%.

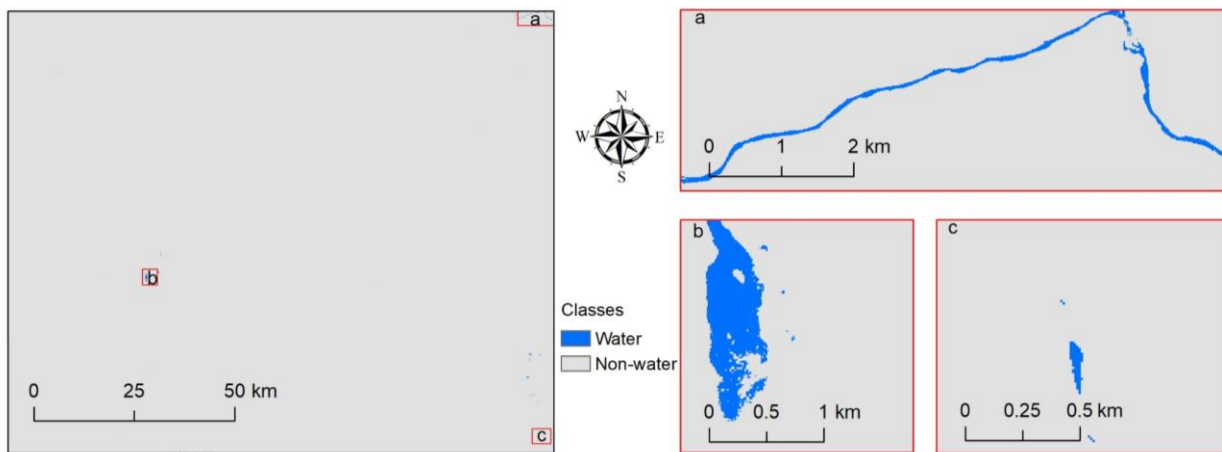


Figure 20 Surface water map of the Amboseli ecosystem during the dry season of 2015 based on Sentinel-1 radar imagery, Sentinel-2 optical imagery and DEM data.

Table 11 The confusion matrix of classification using Sentinel-2 optical and DEM data during the wet season.

Prediction	Reference		User's accuracy
	1	2	
1-Water	36	0	100.00%
2-Non-water	4	40	90.91%
Producer's accuracy	90.00%	100.00%	
	Overall accuracy		95.00%
	Kappa		0.90

Figure 21 shows the classification result in the Amboseli ecosystem during the wet season using Sentinel-2 optical data and DEM. Table 12 shows the confusion matrix of the image classification. Among the 40 reference points for each class collected for accuracy assessment, 34 points of water class were correctly classified, and 40 points of Non-water class were correctly classified. The overall accuracy is 92.50% and the kappa coefficient is 0.85. The producer's accuracy of water and Non-water classes are 85.00% and 100.00%, respectively. The user's accuracy of water and Non-water are 100.00% and 86.96%, respectively. The omission errors of water and Non-water are 15.00% and 0%, the commission errors of water and Non-water class are 0% and 13.04%.

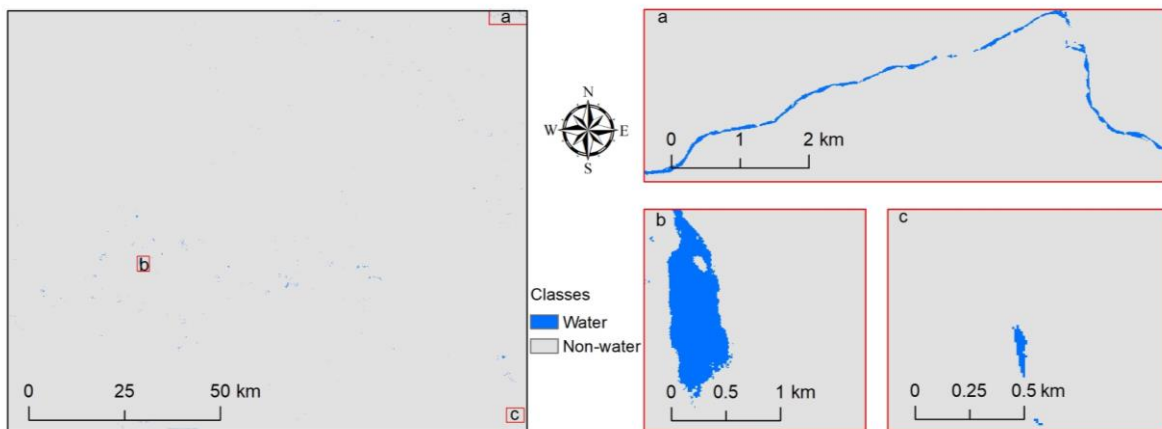


Figure 21 Surface water map of the Amboseli ecosystem during the wet season of 2016 based on Sentinel-1 radar imagery, Sentinel-2 optical imagery and DEM data.

Table 12 The confusion matrix of classification using Sentinel-1, Sentinel-2 and DEM during the wet season.

Prediction	Reference		User's accuracy
	1	2	
1-Water	34	0	100.00%
2-Non-water	6	40	86.96%
Producer's accuracy	85.00%	100.00%	
	Overall accuracy		92.50%
	Kappa		0.85

4.1.4. Accuracy comparison

During the dry season, three random forest models, include model A, B and C, using Sentinel-1 radar, Sentinel-2 optical and the combination of both data was built to map the surface water in the Amboseli ecosystem. Five hundred kappa coefficients for each model were generated. Two-sample t-test ($\alpha = 0.05$) was conducted and results showed that there was a significant difference between the means of kappa coefficients in image classification over the Amboseli ecosystem using Sentinel-1 and the combination of Sentinel-1, Sentinel-2 data in the dry season of 2016 ($p < 0.05$). However, evidence was not enough to reject the null hypothesis that there was no significant difference between the means of kappa coefficients in image classification over the Amboseli ecosystem using Sentinel-2 and the combination of Sentinel-1, Sentinel-2 data in the dry season of 2016 ($p = 0.34$). Thus, surface water map that has the highest kappa coefficient, the map derived from Sentinel-2 optical data was selected for further analysis. Figure 24 shows the error bar of the kappa coefficients.

During the wet season, three random forest models, include model D, E and F, using Sentinel-1 radar, Sentinel-2 optical and the combination of both data was built to map the surface water in the Amboseli ecosystem. Five hundred kappa coefficients for each model were generated. Two-sample t-test ($\alpha = 0.05$) was conducted and results showed that there was a significant difference between the means of kappa coefficients in image classification over the Amboseli ecosystem using Sentinel-1 and the combination of Sentinel-1, Sentinel-2 data in the wet season of 2016. However, evidence was not enough to reject the null hypothesis that there was no significant difference between the means of kappa coefficients in image classification over the Amboseli ecosystem using Sentinel-2 and the combination of Sentinel-1, Sentinel-2 data in the wet season of 2016. Thus, surface water map that has the highest kappa coefficient, the map derived from the combination of Sentinel-1 radar and Sentinel-2 optical data was selected for further analysis.

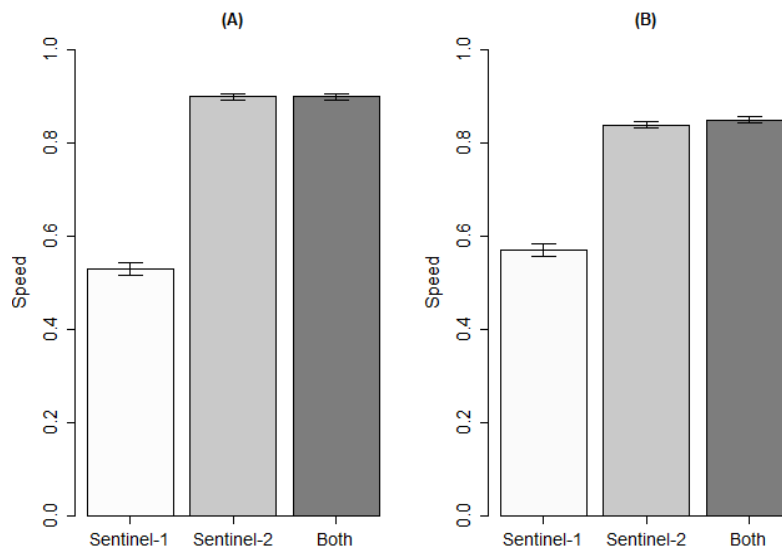


Figure 22 Comparison of kappa coefficients during the dry season (A) and wet season (B).

4.2. Comparison of the surface water

4.2.1. Comparison of the number and area of surface water

As Table 13 shows, the number of rivers captured by Sentinel and JRC (derived from Landsat satellite images) is smaller than that in ILRI surface water product. Sentinel has captured 111 lakes/small water bodies in the dry season. The area of lakes or small water bodies is significantly larger than those in Sentinel and ILRI surface water products. The area of marshes in ILRI water products is substantially larger than Sentinel surface water products.

Table 13 Comparison of the number and area of surface water

Surface water	Rivers	Lakes/small water bodies		Marshes	
	Number	Number	Area (km ²)	Number	Area (km ²)
Sentinel(dry)	1	111	0.73	3	29.9
Sentinel(wet)	1	668	2.35	3	29.9
ILRI	22	1	196.5	6	879.8
JRC (dry)	0	5	0.49	0	0
JRC (wet)	1	245	3.43	0	0

4.2.2. Comparison of the mean distances to surface water

Euclidean distance to the surface water was generated based on different surface water products. Figure 23 shows the distance to surface water in dry and wet seasons calculated from Sentinel water products. Figure 24 shows the distance to surface water in dry and wet seasons calculated from ILRI water products. Figure 25 shows the distance to surface water in dry and wet seasons calculated from JRC water products.

Euclidean distance to surface water in dry and wet season calculated from Sentinel water product

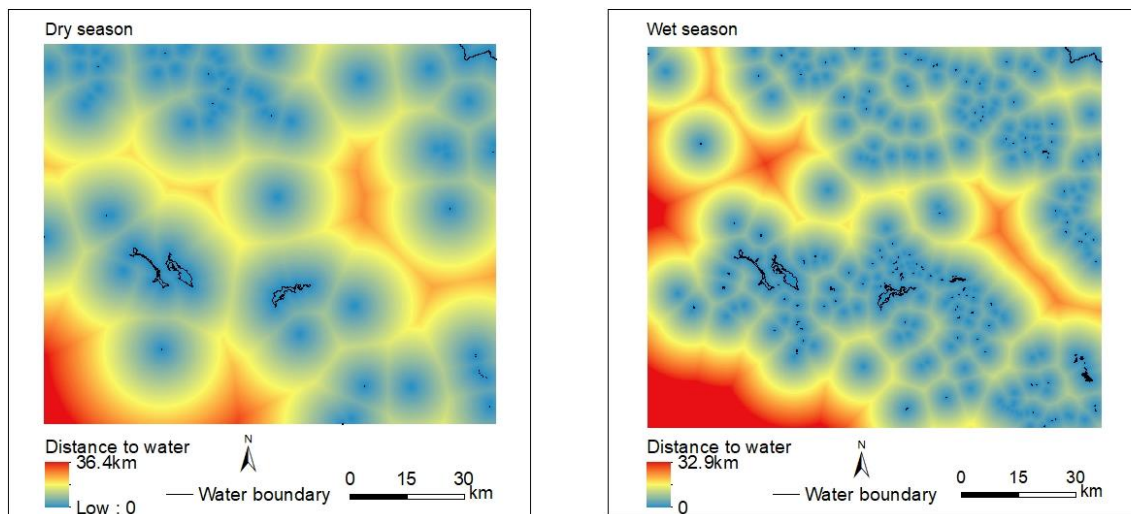


Figure 23 Euclidean distance to surface water in dry and wet season calculated from Sentinel water product.

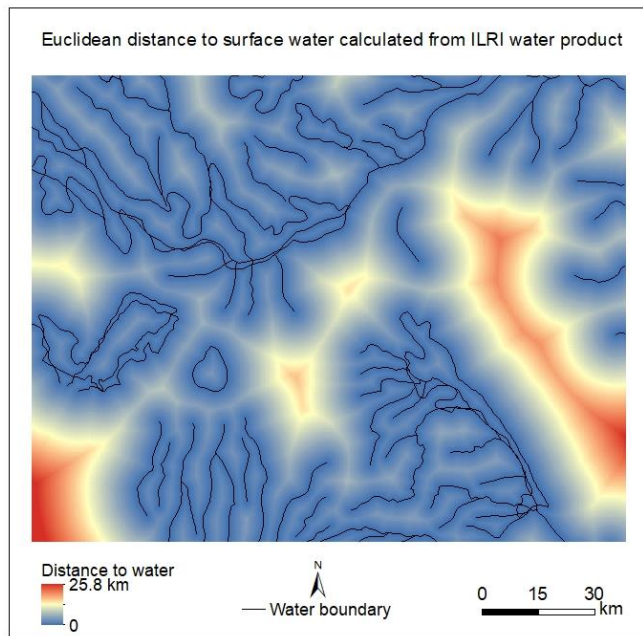


Figure 24 Euclidean distance to surface water in dry and wet season calculated from ILRI water product.

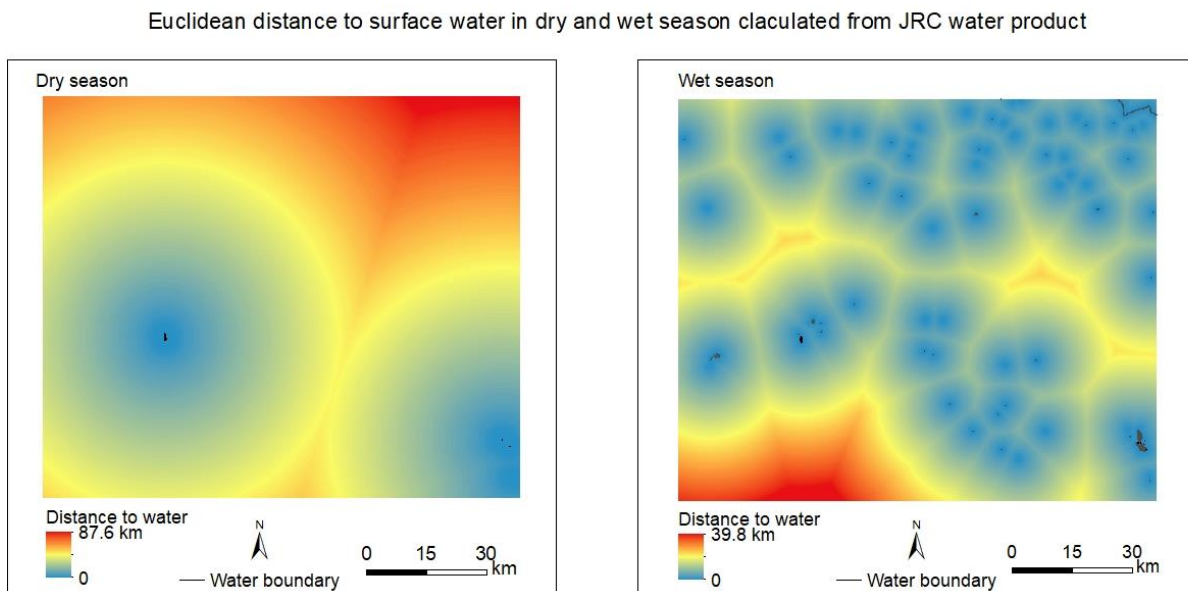


Figure 25 Euclidean distance to surface water in dry and wet season calculated from JRC water product.

Figure 26 shows, during dry and wet seasons, the median of the distance extracted from Sentinel product is larger than the median of the distance derived from ILRI and smaller than the one extracted from the JRC product. The one-way ANOVA test observed that there was a significant difference between the means of the distances (to surface water) derived from the three surface water products during the dry season ($p < 0.05$). A similar result can be observed during the wet season ($p < 0.05$).

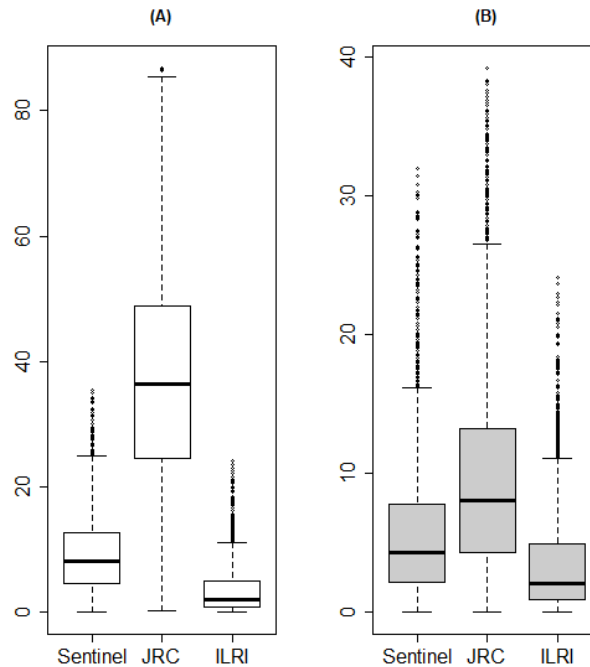


Figure 26 Distances of 3000 random points to the surface water during the dry season (A) and wet season (B). The vertical axis is the distance (km).

4.3. The distance of elephants to the nearest surface water

The mean distances of elephants to the nearest surface water extracted from the Sentinel water products during the dry season and wet season are 6.7 km and 4.2 km, respectively. The mean distances of elephants to the nearest surface water extracted from the ILRI water products during the dry season and wet season are 2.2 km and 2.6 km, respectively. The mean distances of elephants to the nearest surface water extracted from the JRC water products during the dry season and wet season are 29.0 km and 13.2 km, respectively.

The distance of elephants to the nearest surface water in the Amboseli ecosystem during different seasons using different surface water products were tested with Shapiro Wilk (with 0.05 significance level). Results also show that the distance of elephants is not a normal distribution ($p < 0.05$). Mann-Whitney U test results demonstrate that the distance of elephants to the nearest surface water during the dry season is significantly different from the distance during the wet season using surface water derived from Sentinel satellite images, Landsat satellite images. However, there is not any significant difference between the distance of elephants to the nearest surface water during the dry season and wet season using the surface water derived from ILRI ($P > 0.05$). As Figure 22 shows, the distance of elephants to the nearest surface water in the dry season derived from JRC is larger than the distance derived from the other three distances. Moreover, a significant difference was observed in the distance to the nearest surface water using Sentinel derived surface water, Landsat derived surface water maps (Figure 27).

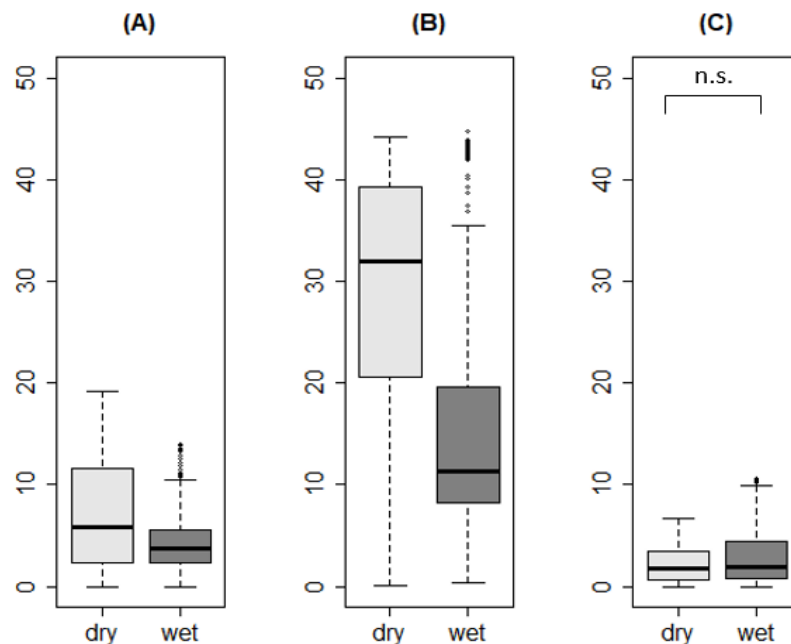


Figure 27 Distance of elephants to the nearest surface water in the Amboseli ecosystem during the dry and wet seasons using Sentinel (A), JRC (B), and ILRI (C) surface water product. The pairs with no significant difference (C) based on the Mann-Whitney U test with $p > 0.05$ are annotated.

4.4. The relationship between speeds of elephant and the surface water availability

As Figure 28 shows, during the dry season, the Mann-Whitney U test of the speeds in different distance to water categories showed that significant difference did not exist in the speeds of elephant movement between the close (<3 km) and medium distance (3-6 km) to the nearest surface water classes ($p=0.74$). Similarly, the Mann-Whitney U test of the speeds between the medium (3-6 km) and far (>6km) distance to the nearest surface water classes also showed that there was no significant difference in the speeds of these two distance classes ($p=0.19$). However, a significant difference did exist in the speeds of the elephant movement between close (<3 km) and far distance (>6 km) categories ($p<0.05$).

As Figure 28 demonstrates, during the wet season, the Mann-Whitney U test of the speeds in different distance to water classes showed that significant difference did not exist in the speeds of elephant movement between the close and medium distance to the nearest surface water classes ($p=0.86$). However, the Mann-Whitney U test of the speeds between the close and far, medium and far distance to the nearest surface water classes demonstrated that there was no significant difference in the speeds of these distance classes ($p<0.05$).

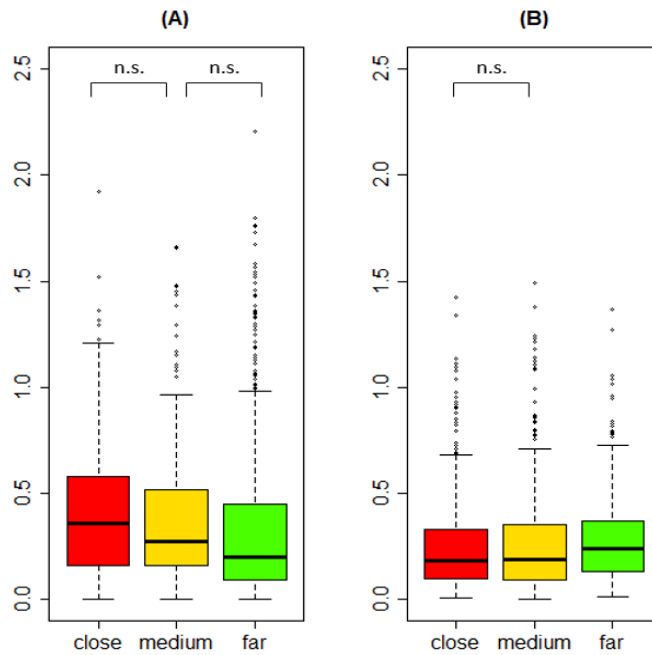


Figure 28 Comparison of speed of elephant movement among three distance to water categories based on surface water derived from Sentinel satellite images. (A) represent the difference in speeds of elephant movement during the dry and (B) represent the one in wet season. The pairs with no significant difference based on the Mann-Whitney U test ($p > 0.05$) are annotated.

Figure 29(A), (B) demonstrate the speed difference between different distances to the nearest surface water classes based on ILRI surface water maps during the dry season. As Figure 29 (A) shows, Mann-Whitney U tests of the speeds in different distance to water classes based on ILRI showed that significant difference did not exist in the speeds of elephant movement between the close and medium distance to the nearest surface water classes ($p=0.36$). Similarly, the Mann-Whitney U test of the speeds between the medium and far distance to the nearest surface water classes also showed that there was no significant difference in the speeds between the medium and far distance classes ($p=0.72$). However, a significant difference did exist in the speeds of the elephant movement between close and far distance classes ($p<0.05$).

As Figure 29 (B) demonstrate, during the wet season, the Mann-Whitney U test of the speeds in different distance to water classes showed that there was not any significant difference in the speeds between the close and medium, medium and far, close and far distance to the nearest surface water classes ($p<0.05$).

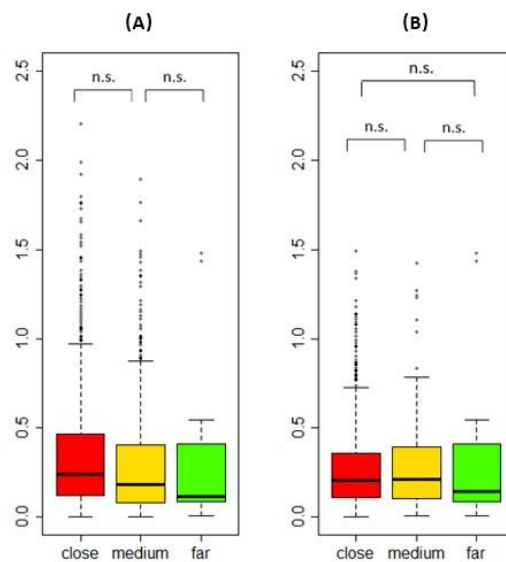


Figure 29 Speeds of elephant movement among three distance to water categories based on surface water produced by ILRI during the dry season(A) and wet season (B). The pairs with no significant difference are annotated based on Mann-Whitney U test with $p > 0.05$.

5. DISCUSSION

5.1. Remote sensing of mapping surface water in the Amboseli ecosystem

The surface water maps over the Amboseli ecosystem have relatively high kappa coefficients despite the fact that the optical images can be easily affected by the cloud, which can be attributed to many factors. First of all, Sentinel-2 images used in this study are with low cloud coverage. Secondly, the cloud and cloud shadow in the Sentinel-2 satellite images are detected using an effective software, Fmask.4.0 since the software had been tested with 67 reference images and the average accuracy is approximately 88% (Qiu et al., 2017). Thirdly, the indices, NDVI and NDWI significantly contribute to high coefficients in mapping the surface water since the variable importance of these two variables are high. The effectiveness of NDWI in delineating surface water has been demonstrated in (McFeeters, 1996). The vegetation NDVI also has been observed to perform well in extracting surface water (Rokni, Ahmad, Selamat, & Hazini, 2014; Zhou et al., 2017). Fourthly, the contribution of DEM in the accuracy of the surface water maps cannot be ignored since the DEM is closely linked with the flow of surface water (Huang et al., 2018).

The kappa coefficients of mapping surface water in the Amboseli ecosystem using Sentinel-1 radar data are relatively small due to many reasons even though Sentinel-1 radar data had been widely and successfully used for water mapping in previous studies. Xing et al. (2018) used Sentinel-1 radar images to map the surface water in Dongting Lake and the kappa coefficients had been observed to be higher than 0.85. Similarly, an automated algorithm was developed to monitor the surface water and the kappa coefficients ranged from 0.879 to 0.910 (Twele, Cao, Plank, & Martinis, 2016). The relatively low kappa coefficients in mapping surface water can be resulted from many reasons. First of all, the water-alike smooth objects, such as pavements and sand surfaces can be wrongly classified as water due to their similarities in the radar backscatter, thus largely increase the commission error in the image classification. Classification accuracies have been observed to be significantly improved by applying the Sand Exclusion Layer (SEL) to map water in arid regions (Martinis, Plank, & Ćwik, 2018). Future study can test the ability to reduce the commission error by applying the extra sand or road masks. Furthermore, the point-based training data may yield lower accuracy compared with polygon-based training data. The influence of point or polygon based training data on the performance of random forest classification has been examined and results showed that polygon based training data in image classification can yield high accuracy (Corcoran, Knight, Pelletier, Rampi, & Wang, 2015). Future work can consider using polygon based instead of point-based training data in the image classification. In addition, the speckle on the radar image affects classification accuracy. The applied Refined Lee speckle filter reduced the resolution of the radar image. A Markov random field (MRF) based filter can be adapted to reduce the speckle without compromising the spatial resolution (Lankoande, Hayat, & Santhanam, 2009). Finally, omission error also can be increased by inundated vegetation and Bragg scattering.

Sentinel surface water maps have been observed to capture a larger number of lakes/small water bodies, especially during the dry season compared with JRC and ILRI surface water products. The difference in the capability of capturing lakes/small water bodies using Sentinel-2 and JRC surface water can be explained by that the selected bands (10 m spatial resolution) of Sentinel-2 satellite images have finer spatial resolution compared with the Landsat satellite images (30 m spatial resolution). Therefore, many lakes/small water bodies can be detected by Sentinel-2 images instead of Landsat images. The advantage of extracting surface water using Sentinel-2 over Landsat satellite images also have been demonstrated in (Kwang, Jnr, & Amoah, 2017). The difference in the capability of capturing lakes/small water bodies using Sentinel-2 and ILRI surface water can be explained by the different mapping scale. The Sentinel maps surface water at a local

scale, while the ILRI maps the surface water at a country scale, thus surface water boundaries may be generalized in the ILRI surface water maps.

5.2. The relationship between the surface water availability and elephant movement in the Amboseli ecosystem

The difference was observed between the distance to the nearest surface water in the dry season and wet seasons using different surface water maps. The distance of elephants to the nearest surface water during the dry season is significantly higher than that of the wet season using surface water derived from Sentinel or Landsat satellite images (JRC). Compared with the vector layer of surface water maps provided by the ILRI, the satellite-derived surface water maps have an obvious advantage over the temporal resolution. Satellite images derived surface water maps have provided surface maps during the dry and wet season. While for the vector-based layer of the surface water map, only one surface water map is available and the season information was missing. Hence, the estimation of the difference between the distance to the nearest surface water in the different season can be biased since not all the surface water are permanent water bodies, especially in a semi-arid landscape. The distance of elephants to the nearest surface water in the dry and wet season showed no significant difference in using the surface water provided by ILRI. This result is not reliable since the surface water maps made by ILRI is on a country scale and the surface water boundary is generalized, the distance of elephants to the nearest surface water is relatively short, which resulted in no difference in distance of elephants to the nearest surface water. The distance of the elephant to the nearest surface water is significantly larger compared with the distance in the dry season derived from other surface water maps

The distance of elephants to the nearest surface water during the dry season has been observed to be larger than that of the wet season in this study. However, the previous study claimed that the distance of elephant to the surface water is shorter compared with those in the wet season (Harris et al., 2008; Park et al., 2002). Elephants are observed to stay close to water during the dry season due to the scarce water (Ashiagbor & Danquah, 2017). This study contradicts with previous studies can be explained in two aspects. On one hand, the adoption of different surface water can explain the part of the difference. Harris et al. (2018) have pointed out that the small and numerous water bodies were not taken into consideration during the wet season and admitted that their research is conservative during the wet season. Park et al. (2002) only considered the perennial water resources while analyzing the distance of elephants to the nearest surface water in their study. On the other hand, even though the elephants prefer to stay closer to water during the dry season does not mean that the elephants are further away from water during the wet season. The distribution and abundance of the surface water during the dry and wet season are different. During the wet season, even if the water does not constrain the movement of the surface water, there are many rain-filled small water bodies (Chamaillé-Jammes et al., 2013), which still can make shorten the elephant's distances to the surface water during the wet season.

During the dry season, a significant difference was observed in the speeds of the elephant movement between close and far distance to the nearest surface water classes ($p < 0.05$). The elephant speeds are relatively larger when they are close to the nearest surface water. As demonstrated before, elephants are water-dependent species. During the dry season, limited surface water resources constrain the elephant movement. This result confirms with previous research. For example, Chamaillé-Jammes et al. (2013) have studied the constraints of surface water on the elephant movement and observed that elephant adjusted speed in response to the constraints, elephants tended to increase the movements speeds when they are arriving or leaving the surface water. Similarly, elephants were observed to adopt high-speed strategy when they are approaching or leaving water resources, this allows the elephants use water for thermoregulation (Thaker, Gupte, Prins, Slotow, & Vanak, 2019).

Time difference between the Sentinel-2 surface water map (2016) and the elephant GPS tracking data (2015) is an issue needed but difficult to address in this study since either elephant GPS tracking data in 2016 or Sentinel-2 satellite images in Sentinel-2 is unfortunately unavailable. Surface water, as well as elephant movement, is dynamic. Two scenarios can be generated in this study. Scenario one is that the surface water distribution in 2015 over the Amboseli ecosystem is similar to that in 2016 based on the assumption that the monthly average rainfall and temperature of these two years are similar. Scenario 2 is that the surface water in 2015 and 2016 are completely different, and this study provided a framework for mapping surface water and analyzing the relationship between the elephant movements and the surface water availability.

Water is widely considered as one of the most important resources around the world. Also, elephants play an essential role in wildlife tourism and local communities can benefit from the prosperity of tourism. Furthermore, the sustainable management of the surface water resources can reduce the human-elephant conflicts to some extent since elephants and livestock compete for limited water resources during the dry seasons in semi-arid landscapes. The derived surface water maps and the findings of the relationship between the elephants and the surface water can assist in developing conservation and management strategies of surface water resources and wildlife resources. Knowing the distribution of the surface water resources is the prerequisites for protecting the resources. Also, the surface water maps and the distances of elephants to the nearest surface water maps can provide information for the allocation of the artificial water points.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

This study mainly aims to quantify the spatiotemporal relationship between elephant movements and surface water availability in the Amboseli ecosystem. And to quantify the relationship, surface water maps during the dry and wet seasons are derived from the Sentinel-1 radar imagery, Sentinel-2 optical imagery and the combination of both. Statistical tests were implemented to test the difference between the kappa coefficients of the surface water maps during the dry and wet seasons using different datasets. Also, the derived Sentinel surface water maps was compared with ILRI and JRC surface water products. Conclusions drawn from the study demonstrate as follows,

First of all, the combination of Sentinel-1 radar data and Sentinel-2 optical data does not always improve the kappa coefficient of surface water mapping compared with the coefficients using either Sentinel-1 or Sentinel-2 data alone. Results of the tests evidence this conclusion. On one hand, the surface water maps during the dry and wet seasons derived from the combination of the Sentinel-1 and Sentinel-2 data yield higher kappa coefficients compared with the ones derived from the Sentinel-1 radar imagery. On the other hands, there is no significant difference between the kappa coefficients of the surface water maps during the dry and wet season derived from the Sentinel-2 optical imagery and the combination of Sentinel-1, Sentinel-2 imagery.

Secondly, the difference between the three surface water products (Sentinel, ILRI and JRC) highlights the advantages of Sentinel surface water products. The difference in water distribution patterns has been detected. The water distribution pattern refers to the number of water bodies or rivers, mean distance to surface water). Sentinel-2 surface water maps capture a large number of small water bodies, especially during the dry season.

Thirdly, elephants are closer to the water during the dry season compared with the wet seasons. The Mann-Whitney U test shows that there is a significant difference between the distances of elephants to the nearest surface water during different seasons based on the surface water products derived from the Sentinel satellite images, Landsat images. During the dry season, surface water is more scarce, while surface water is more abundant during the wet season due to the different rainfalls of these two seasons. Thus the distances between the elephants and the surface water are shorter during the dry season compared with the ones in the wet seasons. No significant difference was observed between the distances of the elephants to the nearest surface water is caused by the relatively abundant surface water in the maps provided by ILRI.

Finally, the movement speeds of the elephant in the Amboseli ecosystem are higher close to surface water compared with further away from surface water during the dry and wet season, respectively. The Mann-Whitney tests results can support this conclusion. Previous research has proved that elephants change speeds in response to surface water constraints.

6.2. Limitations and suggestions for future research

Omission errors in mapping surface water using Sentinel-2 optical data and DEM may mainly be caused by mixed pixels at the water and non-water boundary. In particular, part of the rivers was not detected due to the impurity in pixels caused by vegetation grows along the riverside and narrow width of rivers. A possible approach for addressing the issue of incomplete rivers may be integrating the decomposition of mixed pixels or considering the texture information during the image classification (Yang, Zhao, Qin, Zhao, & Liang, 2017). Commission errors also can be found in mapping surface water using optical satellite images. Built-up area and terrain shadows can be misclassified as surface water, thus led to an overestimation of the surface water. Future study can either replace the NDWI with modified NDWI, which has been claimed to reduce the built-up noise that could be mixed with surface water (Xu, 2006) or to adopt an approach to calibrate the spectra to a normal level as explained in (Kao, Ren, & Lee, 2014). The commission error of mapping surface water using Sentinel-2 satellite imagery and DEM was not detected may be caused by the relatively small size of reference points and purposive selection of sampling strategies for accuracy assessment.

Bias in the accuracy assessment of image classification can be caused by the temporal difference among different satellite images and position errors (reference data). Reference data is the benchmark of assessing accuracy and image classification. Reference data should represent the real ground truth. However, reference data always contains errors and is just another classification (Foody, 2002). The acquisition time difference between satellite images and the finer resolution image for collecting test dataset (reference data) was unavoidable although we tried to make the time as close as possible. The time difference in different satellite images can cause errors in the images interpretation and affect the reliability of the reference points. Furthermore, perfectly in co-registration between the classified and referenced data is the assumption for assessing the accuracy of the maps (Verbyla & Hammond, 1995). However, this is always not the case in reality due to the inherent positional errors of the classified and referenced data and this imperfectly match often cause conservative estimation of the classification errors.

Modified Normalized Difference Water Index (MNDWI) and Pan-sharpening can be adopted for further improving the classification accuracy of the surface water maps. Compared with NDWI, MNDWI had been claimed and observed to increase the water maps since it can better differentiate between the pixels of water and the built-up areas. The MNDWI substitutes the near-infrared band (NIR) with the mid-infrared band (MIR) However, the spatial resolution of MIR bands in Sentinel-2 satellite images is 20 m. Pan-sharpening can be used to increase the spatial resolution of a multispectral band in Sentinel-2 satellite images.

A finer temporal resolution of elephants GPS tracking data is recommended for further understanding and further analysis with elephant movement speeds. Whether finer resolution of elephants GPS tracking data with a higher frequency can yield similar or different results needed to be tested. The four-hour temporal resolution of GPS tracking data may not provide enough details about the movement speeds of the elephant since the elephant can go back to the original location after four hours. Attention also should be paid to the quality of the GPS tracking data. Also, future analysis of the relationship between the elephant movements and the surface water could add more samples of the elephants to make the samples as representative as possible.

Further analysis of the elephant movements also can investigate other movement parameters, include the turning angle and the net squared displacement. The analysis of the different combinations of movement of the speeds and the turning angles can reveal the information about the elephant behavior states. Net squared displacement describes the distances between can provide information about the movement patterns.

LIST OF REFERENCES

- Abdikan, S., Sanli, F. B., Ustuner, M., & Calò, F. (2016). Land cover mapping using sentinel-1 SAR data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41(July), 757–761. <https://doi.org/10.5194/isprsarchives-XLI-B7-757-2016>
- Aldenhoff, W., Heuzé, C., & Eriksson, L. E. B. (2018). Comparison of ice/water classification in Fram Strait from C- and L-band SAR imagery. *Annals of Glaciology*, 59(76), 122–113. <https://doi.org/10.1017/aog.2018.7>
- Aschbacher, J., & Milagro-Pérez, M. P. (2012). The European Earth monitoring (GMES) programme: Status and perspectives. *Remote Sensing of Environment*, 120(2012), 3–8. <https://doi.org/10.1016/j.rse.2011.08.028>
- Ashiagbor, G., & Danquah, E. (2017). Seasonal habitat use by Elephants (*Loxodonta africana*) in the Mole National Park of Ghana. *Ecology and Evolution*, 7(11), 3784–3795. <https://doi.org/10.1002/ece3.2962>
- Balzter, H., Cole, B., Thiel, C., & Schmullius, C. (2015). Mapping CORINE land cover from Sentinel-1A SAR and SRTM digital elevation model data using random forests. *Remote Sensing*, 7(11), 14876–14898. <https://doi.org/10.3390/rs71114876>
- Bergstrom, R., & Skarpe, C. (1999). The abundance of large wild herbivores in a semi-arid savanna in relation to seasons, pan and livestock. *African Journal of Ecology*, 37(1), 26–37. <https://doi.org/10.1046/j.1365-2028.1999.00165.x>
- Blanca, M., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 552–557. <https://doi.org/10.7334/psicothema2016.383>
- Bragg, M., Broeren, A., Addy, H., Potapczuk, M., Guffond, D., & Montreuil, E. (2007). Airfoil Ice-Accretion Aerodynamic Simulation. In *45th ALAA Aerospace Sciences Meeting and Exhibit*. Reston, Virginia: American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/6.2007-85>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
- BurnSilver, S. B., Worden, J., & Boone, R. B. (2008). Processes of fragmentation in the Amboseli ecosystem, Southern Kajiado District, Kenya. *Fragmentation in Semi-Arid and Arid Landscapes: Consequences for Human and Natural Systems*, 225–253. <https://doi.org/10.1007/978-1-4020-4906-4-10>
- Chamaillé-Jammes, S., Mtare, G., Makuwe, E., & Fritz, H. (2013). African Elephants Adjust Speed in Response to Surface-Water Constraint on Foraging during the Dry-Season. *PLoS ONE*, 8(3), e59164. <https://doi.org/10.1371/journal.pone.0059164>
- Chamaillé-Jammes, S., Valeix, M., & Fritz, H. (2007). Managing heterogeneity in elephant distribution: Interactions between elephant population density and surface-water availability. *Journal of Applied Ecology*, 44(3), 625–633. <https://doi.org/10.1111/j.1365-2664.2007.01300.x>
- Clement, M. A., Kilsby, C. G., & Moore, P. (2017). Multi-temporal synthetic aperture radar flood mapping using change detection. *Journal of Flood Risk Management*. <https://doi.org/10.1111/jfr3.12303>
- Clerici, N., Valbuena Calderón, C. A., & Posada, J. M. (2017). Fusion of sentinel-1a and sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia. *Journal of Maps*, 13(2), 718–726. <https://doi.org/10.1080/17445647.2017.1372316>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurements*, 20(1), 37–46.
- Corcoran, J., Knight, J., Pelletier, K., Rampi, L., & Wang, Y. (2015). The effects of point or polygon based training data on randomForest classification accuracy of wetlands. *Remote Sensing*, 7(4), 4002–4025. <https://doi.org/10.3390/rs70404002>
- De Beer, Y., Kilian, W., Versfeld, W., & Van Aarde, R. J. (2006). Elephants and low rainfall alter woody vegetation in Etosha National Park, Namibia. *Journal of Arid Environments*, 64(3), 412–421. <https://doi.org/10.1016/j.jaridenv.2005.06.015>
- de Beer, Y., & van Aarde, R. J. (2008a). Do landscape heterogeneity and water distribution explain aspects of elephant home range in southern Africa's arid savannas? *Journal of Arid Environments*, 72(11), 2017–2025. <https://doi.org/10.1016/j.jaridenv.2008.07.002>
- de Beer, Y., & van Aarde, R. J. (2008b). Do landscape heterogeneity and water distribution explain aspects of elephant home range in southern Africa's arid savannas? *Journal of Arid Environments*, 72(11), 2017–2025. <https://doi.org/10.1016/J.JARIDENV.2008.07.002>
- Dietz, A. J., Klein, I., Gessner, U., Frey, C. M., Kuenzer, C., & Dech, S. (2017). Detection of water bodies from AVHRR data-a TIMELINE thematic processor. *Remote Sensing*, 9(1). <https://doi.org/10.3390/rs9010057>

- Dirzo, R., Hayward, M. W., Kerley, G. I. H., Valkenburgh, B. Van, Ripple, W. J., Newsome, T. M., Sandom, C. J. (2015). Collapse of the world's largest herbivores. *Science Advances*, 1(4), e1400103. <https://doi.org/10.1126/sciadv.1400103>
- Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., & Li, X. (2016). Water bodies' mapping from Sentinel-2 imagery with Modified Normalized Difference Water Index at 10-m spatial resolution produced by sharpening the swir band. *Remote Sensing*, 8(4). <https://doi.org/10.3390/rs8040354>
- Dunkin, R. C., Wilson, D., Way, N., Johnson, K., & Williams, T. M. (2013). Climate influences thermal balance and water use in African and Asian elephants: physiology can predict drivers of elephant distribution. *Journal of Experimental Biology*, 216(15), 2939–2952. <https://doi.org/10.1242/jeb.080218>
- Easterling, D. R., Meehl, G. A., Parmesan, C., Changnon, S. A., Karl, T. R., & Mearns, L. O. (2000). Climate extremes: Observations, modeling, and impacts. *Science*, 289(5487), 2068–2074. <https://doi.org/10.1126/science.289.5487.2068>
- Farr, T. G., & Kobrick, M. (2000). Shuttle radar topography mission produces a wealth of data. *Eos*, 81(48), 583–585. <https://doi.org/10.1029/EO081i048p00583>
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Frantz, D., Haß, E., Uhl, A., Stoffels, J., & Hill, J. (2018). Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*, 215, 471–481. <https://doi.org/10.1016/j.rse.2018.04.046>
- Fuller, A., Mitchell, D., Maloney, S. K., & Hetem, R. S. (2016). Towards a mechanistic understanding of the responses of large terrestrial mammals to heat and aridity associated with climate change. *Climate Change Responses*, 3(1), 10. <https://doi.org/10.1186/s40665-016-0024-1>
- Ganguly, S., & Delhi, N. (2016). Role of Water in Livestock. *The Rec. Adv. Acad. Sci. J.* 1,56-50.
- Gara, T. W., Wang, T., Skidmore, A. K., Zengeya, F. M., Ngene, S. M., Murwira, A., & Ndaimani, H. (2017). Understanding the effect of landscape fragmentation and vegetation productivity on elephant habitat utilization in Amboseli ecosystem, Kenya. *African Journal of Ecology*, 55(3), 259–269. <https://doi.org/10.1111/aje.12346>
- Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, 27(4), 294–300. <https://doi.org/10.1016/j.patrec.2005.08.011>
- Goldstein, B. A., Polley, E. C., & Briggs, F. B. S. (2011). Random forests for genetic association studies. *Statistical Applications in Genetics and Molecular Biology*, 10(1). <https://doi.org/10.2202/1544-6115.1691>
- Guerschman, J. P., Warren, G., Byrne, G., Lymburner, L., Mueller, N., & Dijk, A. Van. (2011). MODIS-based standing water detection for flood and large reservoir mapping: algorithm development and applications for the Australian continent. *Www.Csiro.Au*, (July 2016), 88. Retrieved from <http://www.csiro.au/publications/waterforahealthycountry/wirada/TechReports/WIRADA-MODIS-standing-water.pdf>
- Han, H., Guo, X., & Yu, H. (2017). Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, 219–224. <https://doi.org/10.1109/ICSESS.2016.7883053>
- Harris, G. M., Russell, G. J., Van Aarde, R. I., & Pimm, S. L. (2008). Rules of habitat use by elephants *Loxodonta africana* in southern Africa: Insights for regional management. *Oryx*, 42(1), 66–75. <https://doi.org/10.1017/S0030605308000483>
- Horning, N. (2010). Random Forests: An algorithm for image classification and generation of continuous fields data sets. *International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences 2010*, Hanoi, Vietnam. <https://doi.org/10.5244/C.22.54>
- Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A Review. *Reviews of Geophysics*, 56(2), 333–360. <https://doi.org/10.1029/2018RG000598>
- Huang, W., Devries, B., Huang, C., Lang, M. W., Jones, J. W., Creed, I. F., & Carroll, M. L. (2018). Automated extraction of surface water extent from Sentinel-1 data. *Remote Sensing*, 10, 797. <https://doi.org/10.3390/rs10050797>
- Irwin, K., Braun, A., Fotopoulos, G., Roth, A., & Wessel, B. (2018). Assessing single-polarization and dual-polarization TerraSAR-X data for surface water monitoring. *Remote Sensing*, 10(6). <https://doi.org/10.3390/rs10060949>
- ISLAM, M. M., & SADO, K. (2000). Development of flood hazard maps of Bangladesh using NOAA-AVHRR images with GIS. *Hydrological Sciences Journal*, 45(3), 337–355. <https://doi.org/10.1080/02626660009492334>

- Jain, S. K., Saraf, A. K., Goswami, A., & Ahmad, T. (2006). Flood inundation mapping using NOAA AVHRR data. *Water Resources Management*, 20(6), 949–959. <https://doi.org/10.1007/s11269-006-9016-4>
- Janitza, S., & Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PLoS ONE*, 13, e0201904. <https://doi.org/10.1371/journal.pone.0201904>
- Janitza, S., Tutz, G., & Boulesteix, A.-L. (2014). Random Forests for Ordinal Response Data: Prediction and Variable Selection. *Computational Statistics and Data Analysis*, 96, 57–73. <https://doi.org/10.1016/j.csda.2015.10.005>
- Journal, A. I. (2015). Comparison of Various Similarity Measure Techniques for Generating Recommendations for E-commerce Sites and Social Websites. *American International Journal of Research in Science, Technology, Engineering & Mathematics*, 219–221.
- Kao, H.-M., Ren, H., & Lee, C.-S. (2014). Calibrated ratio approach for vegetation detection in shaded areas. *Journal of Applied Remote Sensing*, 8(1), 083543. <https://doi.org/10.1117/1.JRS.8.083543>
- Kaplan, G., & Avdan, U. (2017). Mapping and monitoring wetlands using SENTINEL-2 satellite imagery. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4, 271–277. <https://doi.org/10.5194/isprs-annals-IV-4-W4-271-2017>
- Kaplan, G., & Avdan, U. (2017). Object-based water body extraction model using Sentinel-2 satellite imagery. *European Journal of Remote Sensing*, 50(1), 137–143. <https://doi.org/10.1080/22797254.2017.1297540>
- Kaptué, A. T., Hanan, N. P., & Prihodko, L. (2013). Characterization of the spatial and temporal variability of surface water in the Soudan-Sahel region of Africa. *Journal of Geophysical Research: Biogeosciences*, 118(4), 1472–1483. <https://doi.org/10.1002/jgrg.20121>
- Kim, J. W., Lu, Z., Jones, J. W., Shum, C. K., Lee, H., & Jia, Y. (2014). Monitoring Everglades freshwater marsh water level using L-band synthetic aperture radar backscatter. *Remote Sensing of Environment*, 150, 66–81. <https://doi.org/10.1016/j.rse.2014.03.031>
- Klapproth, J. C., & Johnson, J. E. (2009). Riparian Forest Buffers : Effects on Plant and Animal Communities Understanding the Science Behind Riparian Forest Buffers : Effects on Plant and Animal Communities. *Forestry*.
- Kulkarni, V. Y. (2013). Efficient Learning of Random Forest Classifier using Disjoint Partitioning Approach. *Proceedings of the World Congress on Engineering*, 2, 1–5.
- Kummu, M., de Moel, H., Ward, P. J., & Varis, O. (2011). How close do we live to water? a global analysis of population distance to freshwater bodies. *PLoS ONE*, 6(6). <https://doi.org/10.1371/journal.pone.0020578>
- Kwang, C., Jnr, E. M. O., & Amoah, A. S. (2017). Comparing of Landsat 8 and Sentinel 2A using Water Extraction Indexes over Volta River. *Journal of Geography and Geology*, 10(1), 1-7. <https://doi.org/10.5539/jgg.v10n1p1>
- Lang, M. W., Kasischke, E. S., Prince, S. D., & Pittman, K. W. (2008). Assessment of C-band synthetic aperture radar data for mapping and monitoring Coastal Plain forested wetlands in the Mid-Atlantic Region, U.S.A. *Remote Sensing of Environment*, 112(11), 4120–4130. <https://doi.org/10.1016/j.rse.2007.08.026>
- Lankoande, O., Hayat, M. M., & Santhanam, B. (2009). A Markov-random-field-based filter for speckle reduction in ultrasound imagery. *Proc. SPIE 7265, Medical Imaging 2009: Ultrasonic Imaging and Signal Processing*, Lake Buena Vista (Orlando Area). <https://doi.org/10.1117/12.811337>
- Lavigne, C., Ricci, B., Franck, P., & Senoussi, R. (2010). Spatial analyses of ecological count data: A density map comparison approach. *Basic and Applied Ecology*, 11(8), 734–742. <https://doi.org/10.1016/j.baae.2010.08.011>
- Li, J., & Roy, D. P. (2017). A global analysis of Sentinel-2a, Sentinel-2b and Landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sensing*, 9(9). <https://doi.org/10.3390/rs9090902>
- Li, S., Sun, D., Goldberg, M., & Stefanidis, A. (2013). Derivation of 30-m-resolution water maps from TERRA/MODIS and SRTM. *Remote Sensing of Environment*, 134, 417–430. <https://doi.org/10.1016/j.rse.2013.03.015>
- Liu, W., Gopal, S., & Woodcock, C. E. (2004). Uncertainty and Confidence in Land Cover Classification Using a Hybrid Classifier Approach. *Photogrammetric Engineering & Remote Sensing*, 70(8), 963–971. <https://doi.org/10.14358/PERS.70.8.963>
- Lopez, C. V., & Anglberger, H. (2017). Fusion of very high resolution SAR and optical images for the monitoring of urban areas, 2017 Joint Urban Remote Sensing Event (JURSE) (pp.1-4), Dubai. <https://doi.org/10.1109/JURSE.2017.7924551>

- Makindi, S. M., Mutinda, M. N., Olekaikai, N. K. W., Olelebo, W. L., & Abdillahi, A. (2014). Human-Wildlife Conflicts : Causes and Mitigation Measures in Tsavo Conservation Area , Kenya. *International Journal of Science and Research*, 3(6), 1025–1031.
- Markert, K. N., Chishtie, F., Anderson, E. R., Saah, D., & Griffin, R. E. (2018). On the merging of optical and SAR satellite imagery for surface water mapping applications. *Results in Physics*, 9, 275–277. <https://doi.org/10.1016/j.rinp.2018.02.054>
- Martinis, S., Kuenzer, C., Wendleder, A., Huth, J., Twele, A., Roth, A., & Dech, S. (2015a). Comparing four operational SAR-based water and flood detection approaches. *International Journal of Remote Sensing*, 36(13), 3519–3543. <https://doi.org/10.1080/01431161.2015.1060647>
- Martinis, S., Kuenzer, C., Wendleder, A., Huth, J., Twele, A., Roth, A., & Dech, S. (2015b). Comparing four operational SAR-based water and flood detection approaches. *International Journal of Remote Sensing*, 36(13), 3519–3543. <https://doi.org/10.1080/01431161.2015.1060647>
- Martinis, S., Plank, S., & Ćwik, K. (2018). The use of Sentinel-1 time-series data to improve flood monitoring in arid areas. *Remote Sensing*, 10(4). <https://doi.org/10.3390/rs10040583>
- Martinis, S., & Rieke, C. (2015). Backscatter analysis using multi-temporal and multi-frequency SAR data in the context of flood mapping at River Saale, Germany. *Remote Sensing*, 7(6), 7732–7752. <https://doi.org/10.3390/rs70607732>
- McFeeters, S. K. (2013). Using the normalized difference water index (ndwi) within a geographic information system to detect swimming pools for mosquito abatement: A practical approach. *Remote Sensing*, 5(7), 3544–3561. <https://doi.org/10.3390/rs5073544>
- McFEETERS, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432. <https://doi.org/10.1080/01431169608948714>
- Michael, Y., Lensky, I. M., Brenner, S., Tchetchik, A., Tessler, N., & Helman, D. (2018). Economic assessment of fire damage to urban forest in the wildland-urban interface using planet satellites constellation images. *Remote Sensing*, 10(9). <https://doi.org/10.3390/rs10091479>
- Morandeira, N. S., Grimson, R., & Kandus, P. (2016). Assessment of SAR speckle filters in the context of object-based image analysis. *Remote Sensing Letters*, 7(2), 150–159. <https://doi.org/10.1080/2150704X.2015.1117153>
- Mose, V. N., Nguyen-Huu, T., Western, D., Auger, P., & Nyandwi, C. (2013). Modelling the dynamics of migrations for large herbivore populations in the Amboseli National Park, Kenya. *Ecological Modelling*, 254, 43–49. <https://doi.org/10.1016/j.ecolmodel.2013.01.016>
- Nath, R. K., & Deb, S. K. (2010). Water-Body Area Extraction From High Resolution Satellite Images-An Introduction, Review, and Comparison. *International Journal of Image Processing (IJIP)*, 3(3), 353–372. Retrieved from <http://www.cscjournals.org/csc/manuscriptinfo.php?ManuscriptCode=67.68.67.74.39.43.42.48.104>
- Ndaimani, H., Murwira, A., Masocha, M., & Zengeya, F. M. (2017). Elephant (*Loxodonta africana*) GPS collar data show multiple peaks of occurrence farther from water sources. *Cogent Environmental Science*, 3(1), 1–11. <https://doi.org/10.1080/23311843.2017.1420364>
- Ngene, S., Njumbi, S., Nzisa, M., Kimitei, K., Mukeka, J., Muya, S., ... Omondi, P. (2013). Status and trends of the elephant population in the Tsavo – Mkomazi ecosystem. *Pachyderm*, 53(53), 38–50.
- Okello, M. M., Kenana, L., Maliti, H., Kiringe, J. W., Kanga, E., Warinwa, F., ... Mwangi, P. (2016). Population density of elephants and other key large herbivores in the Amboseli ecosystem of Kenya in relation to droughts. *Journal of Arid Environments*, 135, 64–74. <https://doi.org/10.1016/j.jaridenv.2016.08.012>
- Omar, H., Misman, M. A., & Kassim, A. R. (2017). Synergetic of PALSAR-2 and Sentinel-1A SAR Polarimetry for Retrieving Aboveground Biomass in Dipterocarp Forest of Malaysia. *Applied Sciences*, 7(7), 675. <https://doi.org/10.3390/app7070675>
- Park, C. N., Stokke, S., & Toit, J. T. (2002). Sexual segregation in habitat use by elephants in. *East*, 40(360), 360–371. <https://doi.org/10.1046/j.1365-2028.2002.00395.x>
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422. <https://doi.org/10.1038/nature20584>
- Purdon, A., & van Aarde, R. J. (2017). Water provisioning in Kruger National Park alters elephant spatial utilisation patterns. *Journal of Arid Environments*, 141, 45–51. <https://doi.org/10.1016/j.jaridenv.2017.01.014>
- Qiu, S., He, B., Zhu, Z., Liao, Z., & Quan, X. (2017). Improving Fmask cloud and cloud shadow detection in mountainous area for Landsats 4–8 images. *Remote Sensing of Environment*, 199(2017), 107–119.

- <https://doi.org/10.1016/j.rse.2017.07.002>
- Ray, N., & Burgman, M. A. (2006). Subjective uncertainties in habitat suitability maps. *Ecological Modelling*, 195(3–4), 172–186. <https://doi.org/10.1016/j.ecolmodel.2005.11.039>
- Redfern, J. V., Grant, R., Biggs, H., & Getz, W. M. (2015). Surface-Water Constraints on Herbivore Foraging in the Kruger National Park. *Ecological Society of America*, 84(8), 2092–2107.
- Rhyman Purnamasayangsukasih, P., Norizah, K., Ismail, A. A. M., & Shamsudin, I. (2016). A review of uses of satellite imagery in monitoring mangrove forests. *IOP Conference Series: Earth and Environmental Science*, 37(1). <https://doi.org/10.1088/1755-1315/37/1/012034>
- Rokni, K., Ahmad, A., Selamat, A., & Hazini, S. (2014). Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sensing*, 6(5), 4173–4189. <https://doi.org/10.3390/rs6054173>
- Rouse, J. W., Hass, R. H., Schell, J. A., & Deering, D. W. (1972). Monitoring Vegetation Systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*.
- Shadrack, N., Moses, M. O., Joseph, M., Shadrack, M., Steve, N., & James, I. (2017). Home range sizes and space use of African elephants (*Loxodonta africana*) in the Southern Kenya and Northern Tanzania borderland landscape. *International Journal of Biodiversity and Conservation*, 9(1), 9–26. <https://doi.org/10.5897/IJBC2016.1033>
- Shannon, G., Matthews, W. S., Page, B. R., Parker, G. E., & Smith, R. J. (2009). The affects of artificial water availability on large herbivore ranging patterns in savanna habitats: a new approach based on modelling elephant path distributions. *Diversity and Distributions*, 15(5), 776–783. <https://doi.org/10.1111/j.1472-4642.2009.00581.x>
- Sharma, R. C., Tateishi, R., Hara, K., & Nguyen, L. V. (2015). Developing Superfine Water Index (SWI) for global water cover mapping using MODIS data. *Remote Sensing*, 7(10), 13807–13841. <https://doi.org/10.3390/rs71013807>
- Sheng, Y., Gong, P., & Xiao, Q. (2001). Quantitative dynamic flood monitoring with NOAA AVHRR. *International Journal of Remote Sensing*, 22(9), 1709–1724. <https://doi.org/10.1080/01431160118481>
- Sibanda, M., Dube, T., Bangamwabo, V. M., Mutanga, O., Shoko, C., & Gumindoga, W. (2016). Understanding the spatial distribution of elephant (*Loxodonta africana*) poaching incidences in the mid-Zambezi Valley, Zimbabwe using Geographic Information Systems and remote sensing. *Geocarto International*, 31(9), 1006–1018. <https://doi.org/10.1080/10106049.2015.1094529>
- Thaker, M., Gupte, P. R., Prins, H. H. T., Slotow, R., & Vanak, A. T. (2019). Fine-Scale Tracking of Ambient Temperature and Movement Reveals Shuttling Behavior of Elephants to Water. *Frontiers in Ecology and Evolution*, 7(4), 1–12. <https://doi.org/10.3389/fevo.2019.00004>
- Twele, A., Cao, W., Plank, S., & Martinis, S. (2016). Sentinel-1-based flood mapping: a fully automated processing chain. *International Journal of Remote Sensing*, 37(13), 2990–3004. <https://doi.org/10.1080/01431161.2016.1192304>
- Vanbelle, S. (2017). Comparing dependent kappa coefficients obtained on multilevel data. *Biometrical Journal*, 59(5), 1016–1034. <https://doi.org/10.1002/bimj.201600093>
- Veena, V., & Kumar, G. (2014). Reconstruction of cloud contaminated information in optical satellite images. 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD), 1-6..
- Verbyla, D. L., & Hammond, T. O. (1995). Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing*, 16(3), 581–587. <https://doi.org/10.1080/01431169508954424>
- Viera, A. J., & Garrett, J. M. (2005). Understanding Interobserver Agreement: The Kappa Statistics. *Family Medicine*, 37(5), 360–363.
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., & Strauss, P. (2018). Sensitivity of Sentinel-1 backscatter to vegetation dynamics: An Austrian case study. *Remote Sensing*, 10(9), 1–19. <https://doi.org/10.3390/rs10091396>
- Wang, H., Yang, F., & Luo, Z. (2016). An experimental study of the intrinsic stability of random forest variable importance measures. *BMC Bioinformatics*, 17(1). <https://doi.org/10.1186/s12859-016-0900-5>
- Wells, W. M., Norbash, A., Ho-Ming Chen, Chung, A. C. S., & Yu, S. C. H. (2003). Multi-modal image registration by minimizing Kullback-Leibler distance between expected and observed joint class histograms., *Proceedings IEEE Computer Society on Computer Vision and Pattern Recognition* (pp. 570-576). Madison, Wisconsin <https://doi.org/10.1109/cvpr.2003.1211518>
- Western, D. (1975). Water availability and its influence on the structure and dynamics of a savannah large

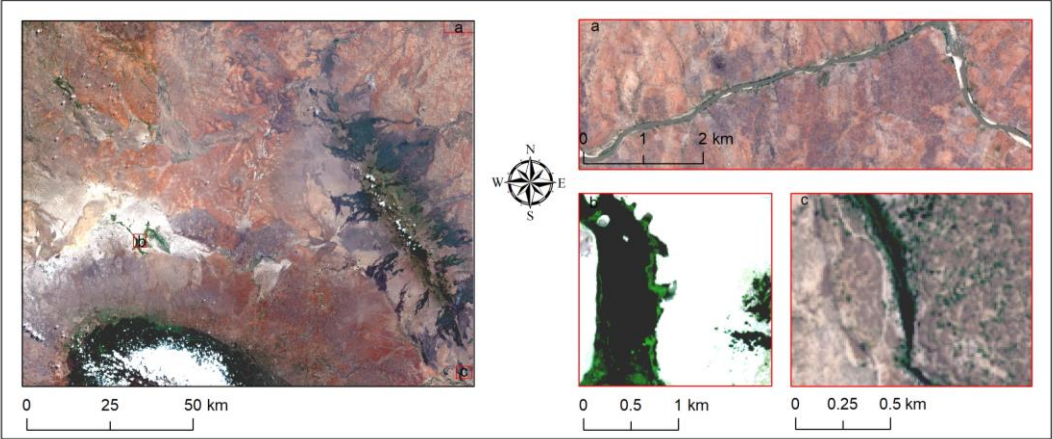
mammal community. *East African Wildlife Journal*, 13, 265–286. <https://doi.org/10.1111/j.1365-2028.1975.tb00139.x>

- Xing, L., Tang, X., Wang, H., Fan, W., & Wang, G. (2018). Monitoring monthly surface water dynamics of Dongting Lake using Sentinel-1 data at 10 m. *PeerJ*, 6, e4992. <https://doi.org/10.7717/peerj.4992>
- Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. <https://doi.org/10.1080/01431160600589179>
- Yang, X., Zhao, S., Qin, X., Zhao, N., & Liang, L. (2017). Mapping of urban surface water bodies from sentinel-2 MSI imagery at 10 m resolution via NDWI-based image sharpening. *Remote Sensing*, 9(6), 1–18. <https://doi.org/10.3390/rs9060596>
- Yano, S., Hanasaki, N., Itsubo, N., & Oki, T. (2015). Water scarcity footprints by considering the differences in water sources. *Sustainability (Switzerland)*, 7(8), 9753–9772. <https://doi.org/10.3390/su7089753>
- Zhou, Y., Dong, J., Xiao, X., Xiao, T., Yang, Z., Zhao, G., ... Qin, Y. (2017). Open surface water mapping algorithms: A comparison of water-related spectral indices and sensors. *Water (Switzerland)*, 9(4). <https://doi.org/10.3390/w9040256>

APPENDICES

Appendix1: True color composite of Sentinel-2 image acquired on July 23, 2016 (A) and December 20, 2016.

(A)



(B)

