

# HYDROLOGICAL & VEGETATION DYNAMICS IN LAKE NGAMI, OKAVANGO DELTA



## **Master's thesis**

Requirement for the degree in

European Master in Environmental Science – Soil, Water and Biodiversity

University of Hohenheim and University of Natural Resources and Life Sciences,  
Vienna

Author: **Eduardo Garcia Bendito**

Supervisor: **Assoc. Prof. Dr. Matthias Bernhardt**. Institute of Hydrology and Water Management, University of Natural Resources and Life Sciences, Vienna, Austria.

Co-supervisor: **Prof. Dr. Frank Schurr**. Institute of Landscape and Plant Ecology, University of Hohenheim, Germany.

Co-supervisor: **Dr. Mathew Herrnegger**. Institute of Hydrology and Water Management, University of Natural Resources and Life Sciences, Vienna, Austria.

Co-supervisor: **Dr. Jörn Pagel**. Institute of Landscape and Plant Ecology, University of Hohenheim, Germany



## **Declaration**

I give my solemn word that I have compiled this work solely and without external help, have not utilized any sources outside those permitted and that the sources used have been given verbatim or quoted textually in the places indicated.



## **Dedication**

To my parents, Carmen and Eduardo, who gifted me the moral codes which accompany me every day of my life. To my sisters, Ana and Martina the true representation of those principles. To Luo, with whom I share those values.

## **Acknowledgements**

First, I would like to thank my supervisors Prof. Dr. Mathias Bernhardt and Dr. Mathew Herrnegger, Dr. Jörn Pagel and Prof. Dr. Frank Schurr. Specially, I would like to thank Matthias and Mathew for their interest in the topic, their invaluable guidance, and their support to gain experiences beyond those purely academic.

Secondly, I would like to thank the people I met along the way of this masters. Specially, I would like to thank Peter, who welcomed me in Vienna and showed me an immense kindness. Also thanks to Feli, who helped me in Stuttgart with her knowledge and friendship. I would also like to thank Sami for his openness and respect.

Last, but not least, I would like to thank Carmen, Eduardo, Ana and Luo for their unconditional love and support through it all.



## Abstract

Hydrology processes are key for vegetation in arid and semi-arid climates, playing an important role in the understanding biogeochemical exchanges between land surface and the atmosphere. Inter-annual variations of hydrological conditions determine water availability for vegetation. The study focuses on the coupling of vegetation and hydrological patterns to assess eco-hydrological interactions in Lake Ngami (Northern Botswana). Earth Observation systems provide consistent and systematic measurements in cases where direct observations and monitoring of environmental variables are not available. Measured surface reflectance may be transformed into more meaningful indicators of environmental conditions. Increasing operability of these systems generate large amounts of data to monitor the environment. Google Earth Engine combines remote sensing data storage and tools to process and analyze a variety of information from these sources. The objective of this study is to evaluate the suitability of this platform to provide reproducible and reliable geo-processing workflows to study the spatio-temporal relationships between vegetation and surface hydrology around Lake Ngami, a sub-basin of the Okavango Delta, from 2000 to 2012. Local hydrological data sets were compared with remote sensing products to evaluate accuracies and discrepancies. Discharge data was compared with the Global Surface Water (GSW) data set, showing a good correlation between the two. Meanwhile near-local observations of precipitation were contrasted with satellite derived site-measurements from the Climate Hazards InfraRed Precipitation with Stations (CHIRPS) data set, showing average differences of 8% between local and satellite measurements. Remote sensing data of surface hydrology and vegetation (MODIS Enhanced Vegetation Index) were used to study eco-hydrological dynamics around Lake Ngami. Results indicate that local phenology is well synchronized with the annual precipitation patterns, while flooding of the lake does not seem to sustain vegetation in the surrounding of Lake Ngami. Overall, it is concluded that a platform such as Google Earth Engine may represent an accessible and reliable platform to access environmental data and monitor ecosystem dynamics in areas lacking long-term monitoring, allowing to identify key ecological processes.

Keywords:

*Eco-hydrology; Remote Sensing; Google Earth Engine; Discharge; Precipitation; Phenology*





## **Abstrakt**

Hydrologische Prozesse sind der Schlüssel für die Vegetation in ariden und semiariden Klimazonen und spielen eine wichtige Rolle beim Verständnis des biogeochemischen Austauschs zwischen Landoberfläche und Atmosphäre. Jährliche Schwankungen der hydrologischen Bedingungen bestimmen die Verfügbarkeit von Wasser für die Vegetation. Die Studie konzentriert sich auf die Kopplung von Vegetation und hydrologischen Mustern, um die öko-hydrologischen Wechselwirkungen im See Ngami (Nordbotswana) zu bewerten. Erdbeobachtungssysteme liefern konsistente und systematische Messungen in Fällen, in denen direkte Beobachtungen und Überwachung von Umgebungsvariablen nicht verfügbar sind. Die gemessene Oberflächenreflexion kann in aussagekräftigere Indikatoren für Umgebungsbedingungen umgewandelt werden. Die Google Earth Engine kombiniert Datenspeicherung und Tools zur Fernerkundung, um eine Vielzahl von Informationen aus diesen Quellen zu verarbeiten und zu analysieren. Das Ziel dieser Studie ist es, die Eignung dieser Plattform für reproduzierbare und zuverlässige Geoverarbeitungs Arbeitsabläufe zu untersuchen, um die räumlich-zeitlichen Beziehungen zwischen Vegetation und Oberflächenhydrologie um den See Ngami, einem Unterbecken des Okavango-Deltas, zwischen 2000 bis 2012 zu erforschen. Die Abflussdaten wurden mit dem Global Surface Water (GSW) -Datensatz verglichen, was eine gute Korrelation zwischen den beiden zeigte. In der Nähe lokaler Niederschlagsbeobachtungen wurden satellitengestützte Standortmessungen aus dem CHIRPS-Datensatz (Climate Hazards InfraRed Precipitation with Stations) gegenübergestellt, was durchschnittliche Unterschiede zwischen lokalen und Satellitenmessungen von 8% zeigten. Fernerkundungsdaten der Oberflächenhydrologie und -vegetation (MODIS Enhanced Vegetation Index) wurden verwendet, um die ökohydrologische Dynamik am See Ngami zu untersuchen. Die Ergebnisse zeigen, dass die lokale Phänologie gut mit den jährlichen Niederschlagsmustern synchronisiert ist, während die Überschwemmungen des Sees die Vegetation in der Umgebung des Ngami-Sees nicht zu unterstützen scheinen. Insgesamt wird die Schlussfolgerung der gezogen, dass eine Plattform wie die Google Earth Engine eine zugängliche und zuverlässige Plattform für den Zugriff auf Umweltdaten und die Überwachung der Ökosystemdynamik in Bereichen sein kann, in denen eine langfristige Überwachung fehlt, wodurch wichtige ökologische Prozesse identifiziert werden können.



# TABLE OF CONTENTS

<b>1. INTRODUCTION.....</b>	<b>1</b>
1.1. PHENOLOGY AND SURFACE HYDROLOGY .....	1
1.2. ENVIRONMENTAL MONITORING AND REMOTE SENSING.....	3
1.3. GOOGLE EARTH ENGINE .....	4
1.4. RESEARCH OBJECTIVES .....	5
<b>2. MATERIALS AND METHODS.....</b>	<b>7</b>
2.1. STUDY AREA .....	7
2.2. MATERIALS.....	13
2.2.1. Local data sets.....	14
2.2.2. Remote Sensing of vegetation .....	17
2.2.3. Remote Sensing of surface water.....	21
2.2.4. Remote Sensing of precipitation.....	22
2.3. METHODOLOGY .....	23
<b>3. RESULTS AND DISCUSSION .....</b>	<b>27</b>
3.1. LOCAL VS. REMOTE SENSING DATA SETS .....	27
3.2. SURFACE HYDROLOGY DYNAMICS .....	31
3.3. VEGETATION DYNAMICS .....	34
3.4. COUPLING OF HYDROLOGICAL AND VEGETATION DYNAMICS.....	37
<b>4. CONCLUSIONS.....</b>	<b>43</b>
<b>5. REFERENCES.....</b>	<b>45</b>
<b>APPENDIX I: GOOGLE EARTH ENGINE SCRIPTS .....</b>	<b>51</b>
APPENDIX I.I: GLOBAL SURFACE WATER (GSW).....	51
APPENDIX I.II: CLIMATE HAZARDS INFRARED PRECIPITATION WITH STATIONS (CHIRPS) .....	55
APPENDIX I.III: MODIS 16-DAY COMPOSITE EVI.....	57
<b>APPENDIX II: GLOBAL SURFACE WATER UNCERTAINTY.....</b>	<b>59</b>
<b>APPENDIX III: ANNUAL PHENOLOGY AND PRECIPITATION (2000 - 2012).....</b>	<b>61</b>





## List of Figures

Figure 1: Map of the Okavango Delta and surroundings (Milzow et al., 2009).....	7
Figure 2: References on the Web Of Science for the study area (2000 - 2017). .....	7
Figure 3: Map of the study area around Lake Ngami. ....	8
Figure 4: Photos of vegetation around Lake Ngami. ....	10
Figure 5: Discharge time-series at Kunyere river (1973 - 2013).....	14
Figure 6: Precipitation time-series at Maun airport (1921 - 2012).....	15
Figure 7: Spectral reflectance curves for vegetation, soil and water (Lillesand et al., 2014).....	17
Figure 8: MODIS vs Landsat.....	19
Figure 9: Google Earth Engine IDE web-interface (Google Earth Engine, 2019) .....	22
Figure 10: Schematic workflow of the methods applied .....	23
Figure 11: Comparison of the discharge vs. lake extent (2000 - 2012).....	25
Figure 12: Comparison of the precipitation data sets (2000 -2012) .....	28
Figure 13: Monthly observations of precipitation and lake extent (2000 - 2012) .....	29
Figure 14: Annual transitions of open water in Lake Ngami (2000 - 2012).....	31
Figure 15: Vegetation dynamics around Lake Ngami (2000 - 2012). ....	33
Figure 16: Length of the rain and growing seasons (2006) .....	35
Figure 17: Minimum and maximum EVI compared with monthly the precipitation in Lake Ngami (2000 - 2012).....	36
Figure 18: Relationship between annual precipitation and maximum EVI each year (2000 - 2012). ....	37
Figure 19: Comparison of monthly EVI response to monthly precipitation and lake extent for each buffer around lake Ngami. ....	38







# 1. Introduction

Water largely determines many dynamics of terrestrial ecosystems directly, regulating physiological processes of vegetation, and indirectly, as a geomorphological driver of abiotic processes on the Earth's surface. Vegetation, on the other hand, influences several hydrological processes such as infiltration and evapotranspiration. Thus, water and vegetation interactions are fundamental processes to understand the productivity and dynamics of ecosystems. As a result, the response of vegetation and its interaction with hydrological components in a context of a changing climate has been discussed and studied widely in the past decades, from global, regional and local scales (White et al., 1997; Choler et al., 2011; Jin et al. 2013; Ma et al. 2013). These interactions are of utmost importance in ecosystem simulations and soil-plant-atmosphere models (White et al., 1997). Furthermore, interactions between vegetation and the surface hydrology play a crucial role in the connectivity of the landscape, on which many ecosystem processes depend (D'Odorico et al., 2010). From precipitation to infiltration, flooding and evapotranspiration, vegetation and water interactions are of major importance to manage water and land resources in many regions of the world, which are increasingly affected by the impacts of changing rainfall and run-off patterns due to climate change.

## 1.1. Phenology and surface hydrology

Ecological processes are to a large extent dominated by hydrological conditions. Particularly vegetation is widely used as a proxy indicator of ecosystem status, based on its potential to provide information on the ecological conditions through spatio-temporal changes and dynamics in the vegetative cover. Phenology (or the annual life cycle of vegetation) is particularly sensitive to small changes in the environment and respond to varying conditions by adjusting the length the photosynthetic season or greening period. Most studies of phenology dynamics have focused on temperature-driven phenology in temperate regions, with small emphasis put on the study of savanna ecosystems (White et al., 1997; Jin et al., 2013), where water is the main limiting factor for organisms and most vegetation dynamics and processes are governed by the presence or lack of water. Under these conditions, water plays a central role in plant physiological processes due to relatively large amounts required during the seasonal plant growth. Similarly, discrete precipitation events temporarily provide enough moisture in the soil to boost vegetative growth (Schwinning et

al., 2004). This is largely the case in the semi-arid savanna of Northern Botswana, where seasonal flooding pulses, rainfall events and distance to surface water bodies may play an important role in controlling the development and extent of vegetation cover (Murray-Hudson et al., 2015). Therefore, landscape dynamics of vegetation communities depend on the periodic presence of water and its availability, particularly in arid and semi-arid regions, where soil water availability is reduced for long periods of time (sometimes for years).

Moreover, many of the inter-relationships between the hydrology and biotic processes (for example, flood-plain ecology) remain poorly understood in semi-arid regions. Research in these fields is associated with areas of interdisciplinary studies and is particularly interesting in new paradigms of water management (Rodriguez-Iturbe, 2000). Most of the initial research in the interface between ecology and hydrology has been carried for hydrological processes in wetlands (Ingram, 1987; Zalewski, 1997) focusing on specific relationships between soil moisture and plants. More recently, these vegetation and hydrology interactions have been receiving growing interest and extended to broader ecosystem definitions (Rodriguez-Iturbe, 2000; D'Odorico et al., 2010). Combination of quantitative data in ecological and hydrological frameworks allows to study the responses and dynamics of vegetation in relation to surface hydrology (Choler et al., 2011). Photosynthetic activity in the canopy of vegetation may be assumed to be a proxy of seasonal dynamics. Phenological variations driven by plant water availability, are key elements to improve our understanding of atmosphere and land surface exchanges and affecting the local water balance due to increased evapotranspiration during the growing season (White et al., 1997; Ge et al., 2016). For this reason, developing methods to obtain key parameters of vegetation dynamics and its response to hydrological environmental drivers could benefit the prediction of regional models, allowing natural resource managers and policy makers to mitigate and adapt to the impacts of landscape changes on local communities.

## 1.2. Environmental monitoring and remote sensing

In general, monitoring of the environment requires long term and consistent measurements of variables that allow representative analyses in the system of interest, with a correct selection of variables to measure and at a relevant scale and appropriate frequency. In-situ monitoring networks often become expensive and in isolated and inaccessible areas such as the Okavango Delta and its surroundings the costs associated with a consistent monitoring network can increase considerably. Moreover, manual monitoring of certain environmental variables lacks continuous sampling over geographic areas. The rise of Earth observation systems in the last decades provides valuable information for monitoring environmental conditions in remote locations through consistent measurements of Earth surface reflectance at different spatio-temporal scales and across entire landscapes. These systems provide a wealth of information on ecosystem dynamics at different scales (Rodriguez-Iturbe, 2000; Mishra et al., 2012), tremendously valuable in regions such as the North of Botswana, where monitoring networks do not exist, or their implementation is not feasible economically (López et al., 2001).

Satellite and remote sensing measurements are based on the reflectance or emission of radiation from different bodies, with the possibility to identify reflectance signatures from specific features such as clouds, water, vegetation, etc. The increase in earth observation has resulted in a large amount of data from different satellite missions equipped with different sensors, which record specific ranges of the electromagnetic spectrum depending on the sensors carried in the mission/platform. Observations from Earth-orbiting satellites are currently one of the most important sources of information for environmental monitoring and research, providing high-resolution measurements both spatially and temporally at different scales, covering virtually the entire surface of the Earth. Recently, and with growing inertia, this information is becoming open and freely accessible from several agencies and organizations including NASA, USGS or the European Space Agency, greatly reducing the costs of environmental monitoring and providing free access to large repositories of data. Data from multi-spectral satellite sensors of missions such as Landsat or MODIS (NASA, USGS) have been translated into a variety of products including measurements on vegetation (Huete et al., 2002), precipitation (Huffman et al., 2001; Funk et al., 2015) and surface water conditions (Yamazaki et al., 2015; Pekel et al., 2016). However, and despite the general agreement on the validity of remote sensing for quantifying ecosystem

dynamics and its ability to provide consistent and repeatable measurements (Mishra et al., 2012), it remains necessary to contrast results from satellite sensors against local measurements.

### 1.3. Google Earth Engine

Reflectance data collected from satellite sensors are generally complex and require special and complex pre-processing techniques in order to correct for different effects and interactions between the radiation and Earth's surface. Image corrections of raw satellite data require expertise as well as enough computation power depending on the amount of data to be handled. Furthermore, storage and size of these data sets are usually measured in terabytes, with high associated costs. With the mentioned increase in free and open Earth Observation, these data sets have become scattered across several portals or repositories, requiring an effort of data mining from the typical basic user.

In 2010 Google Inc. began the development of a platform focused on forest monitoring, which later expanded to a variety of uses related to Earth monitoring, offering cloud computation for Earth Observation products. The platform now includes a repository of well-documented open datasets as well as a large variety of tools to handle and interact with the data. At its current stage, Google Earth Engine represents a one-storage platform for satellite data processing, with a large and rich library of open data products from different satellite missions at any given scale: from local to global. The platform also includes a wide variety of analysis products generated from different satellite missions., as well as secondary thematic data sets on bio-geophysical variables for ecosystem monitoring made available by research groups and scientists. Integration of these thematic datasets allows studying the ecosystem structure and its dynamics, without the necessity to reproduce the entire production phase of the data. Integrating the data, the tools and the infrastructure, Google Earth Engine provides the services to operate a variety of remote sensing products. For these reasons, since its conception, many studies have made use of this platform to conduct research on hydrology, ecology, meteorology, and Earth sciences in general (Gorelick et al., 2017).

#### 1.4. Research objectives

This thesis project was motivated by a previous study at the Okavango Research Institute in Maun (Botswana) in 2015. Observing the differences in presence and health of vegetation depending on the distance to surface water bodies and between dry and wet seasons around Lake Ngami, it was hypothesized that spatial and temporal dynamics of water dominate the local vegetation. However, the lack of quantitative measurements of vegetation and hydrological variables in this area constituted an initial limitation to study this relationship. The current increase of Earth observation systems as well as growing access to data sets and analysis tools, represent a perfect opportunity to study the water-vegetation dynamics around Lake Ngami.

The aim of this study is two-fold: (i) evaluate the use and suitability of the Google Earth Engine platform and remote sensing products as a sampling method in regions where quantitative local measurements of these variables are not available and (ii) investigate the dynamics of vegetation in the surroundings of the Lake Ngami and its relationship to the surface hydrological variables (precipitation and the extent of surface water in the lake). In order to tackle these objectives, two main research questions are proposed to define the theoretical framework of this study:

(1). Are vegetation dynamics in the vicinity of lake Ngami dominated by surface water dynamics and availability?

(2). Does Google Earth Engine represent a viable solution as a storage and processing platform to use remote sensing products and monitor hydrology and vegetation in data scarce regions?



## 2. Materials and methods

### 2.1. Study area

The Okavango River forms in the highlands of Angola and flows southwards through the Caprivi Strip in Namibia, before it spreads into an area of swamps in Northwest Botswana. This endorheic basin is located in Southern Africa and may be considered as an independent system from the Zambezi river (Kling et al., 2014). The river drains into the Okavango wetlands, more well known as the Okavango Delta, a UNESCO World Heritage Site declared in 2012 (UNESCO, 1971). The wetlands spread on an alluvial fan in the Northern Kalahari Desert, creating an area with a mixture of river channels, flood plains and savannas. Local climate change induces higher variability of precipitation in the Okavango Delta, already strongly dominated by inter-annual variable run-off originated in the upper part of the catchment in Angola. Given its location in the North of the Kalahari Desert, the Okavango Delta represent the only perennial surface water body for hundreds of kilometers. The mean extent of the wetlands is about 5,000 km<sup>2</sup>, but intermittently it expands to a maximum of 12,000 km<sup>2</sup> (Figure 1). It is the local hydrological heterogeneity which makes this ecosystem so unique, with a high variance of ecological landscapes in a relative small scale, supporting a large biodiversity and growing tourism industry, although economic development in the form of irrigation systems and high pressure from livestock, as well as global and regional climate change threat this ecosystem included in the RAMSAR List of Wetlands of International Importance (UNESCO, 1971; Mbaiwa, 2005; Milzow et al., 2009).

The Okavango river does not empty in the delta, and part of the streamflow reaching the Southern end is collected in a fault system through a series of channels (Xudum River, Kunyere River, Thaoge River). These channels converge in the Thamalakane and Nhabe rivers, flowing into a system of sub-basins formed by the Makgadikgadi pans in the South-East, the Mababe depression in the North-East and Lake Ngami in the South-West (Figure 1). Possibly due to their vicinity to a thriving enclave such as the Okavango wetlands, the seasonally flooded pans in the Southern fringes have not received much attention (Figure 2; e.g.: Lake Ngami). These areas are of great importance as interfaces and buffer systems between the wetlands and the Kalahari Desert in the South, where transitions between high and low flows have notable impacts on the ecosystem and socio-economic activities in the region.

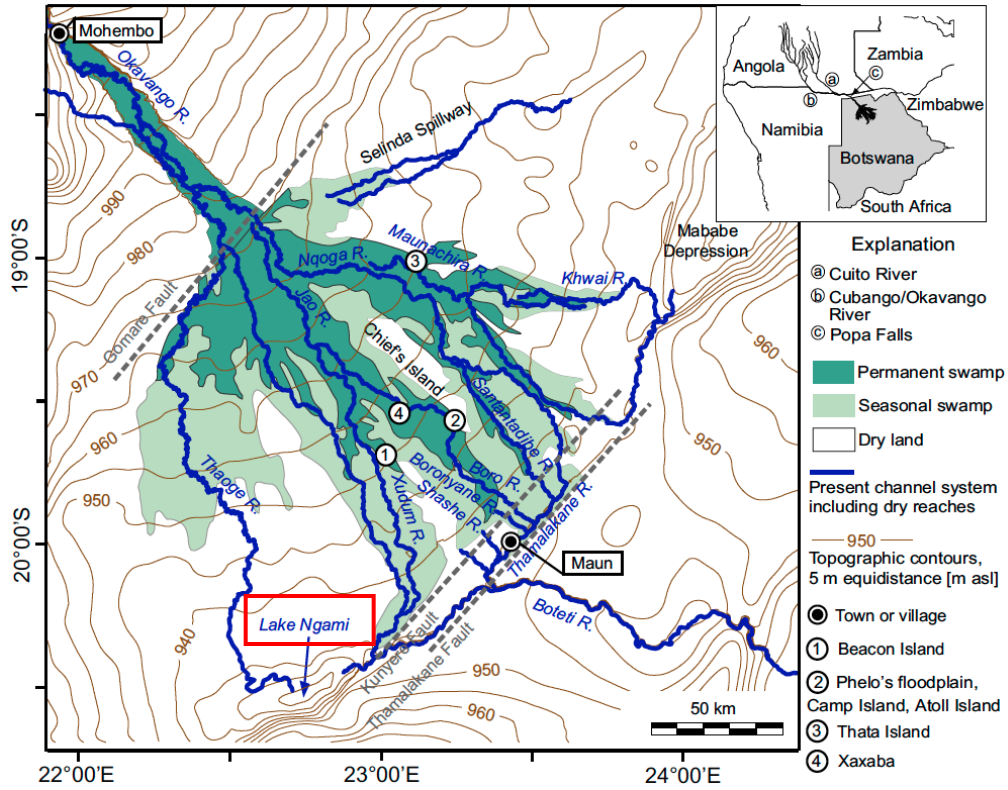


Figure 1: Map of the Okavango basin, delta and its surroundings (Modified after Milzow et al., 2009)

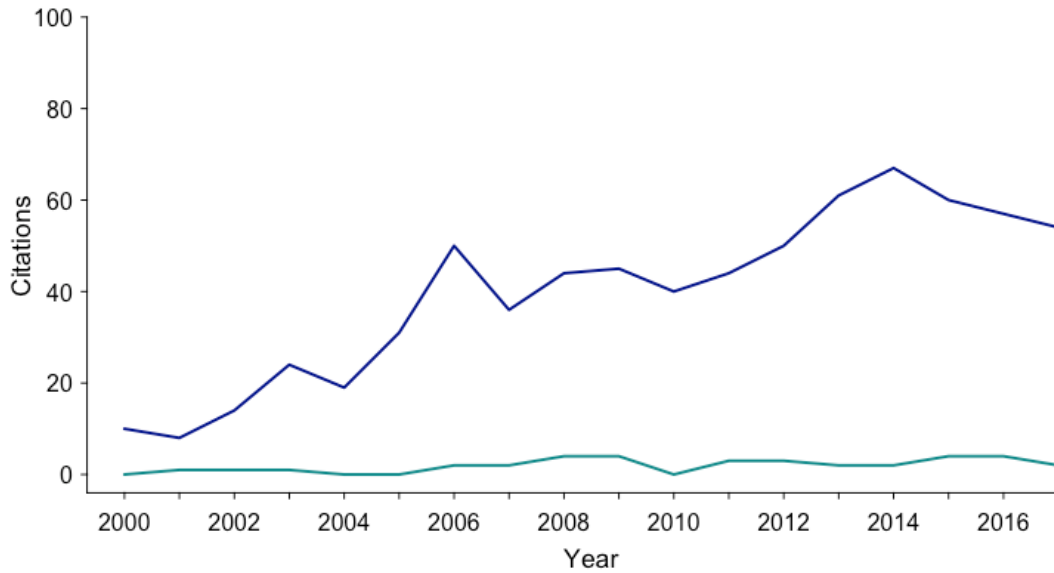


Figure 2: Number of citations on the Web of Science for the terms 'Okavango' (blue) and 'Lake Ngami' (green) in each year between 2000 and 2017.



The study area of interest is located in the southern end of the Okavango region,  $20^{\circ} 30' 7.21''$  S,  $22^{\circ} 44' 46.76''$  E covering ca.  $1500 \text{ Km}^2$ , and includes the recent range of open water in Lake Ngami (since 2000) as well as the surrounding environments (Figure 3). Of the total study area,  $250 \text{ km}^2$  correspond to open water during high discharge years. Lake Ngami is a significant drainage basin in the lower Okavango region, which comprises the area collecting the outflows of the upstream Okavango wetlands (Figure 1).

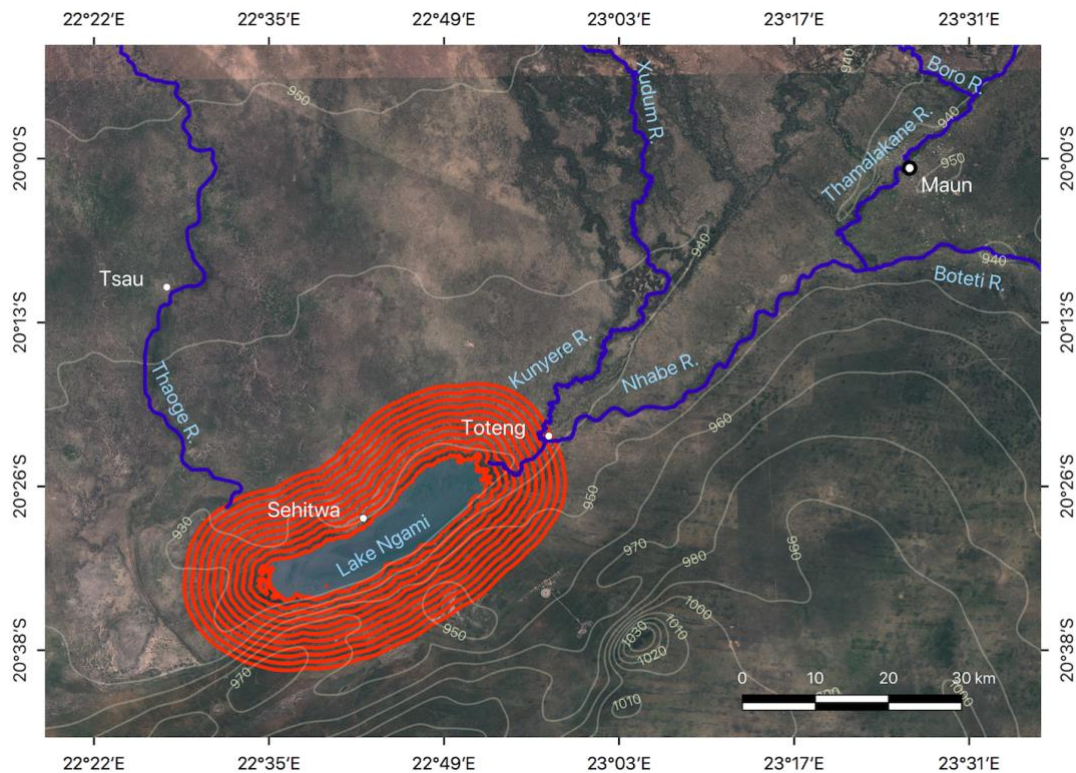


Figure 3: Map of Lake Ngami study area with the 10 buffer stripes of 1 km width (red), which compose the study area. The figure includes the main hydrography of the region (highlighted in blue), contour lines for elevation, as well as the locations of the most relevant villages.

The lake is fed by the Kunyere and Nhabe rivers. These two channels join just upstream of the inlet into the lake at Toteng, although in recent years Nhabe river has not carried water. Similarly, until the 1880s, another channel, the Thaoge river, supplied streamflow into the lake, but ever since this channel has ceased to reach the lake (Kurugundla et al., 2018) resulting in a severe drought and sustained desiccation of Lake Ngami between 1989 and 2004 (Wolski & Murray-Hudson, 2006). The lake is an endorheic, shallow depression, ranging below  $920 \text{ m} - 940 \text{ m}$

above mean sea level (msl), where the limits of the historical lake extension are marked by concentric ridges, each one raising between 2 or 3 meters above the lakebed central pan (Akanyang, 1997). These geological traces seem to indicate the presence of a large inland water body in the Thamalakane Fault due to a wet period about 2000 BP (Shaw, 1985). The lake basin is the result of the Thamalakane and Kunyere fault system dipping South-West, and given the flat topography of the area, surface water flows are very susceptible to geomorphological variations resulting in variations in run-off reaching the lakebed. Flat conditions together with the geologic conformation of the lake, which sits on beds of Kalahari sands, result in a very permeable and shallow aquifer with a variable water table no deeper than a few meters during wet periods (Hamandawana, 2008; Milzow et al., 2009).

The prevailing climatic conditions in Northern Botswana, and particularly in the lake region are semi-arid, with an average annual precipitation of about 450 mm and mean temperatures of 20 °C (Milzow et al., 2009). Temperatures have a typical range of about 20 °C with maximums in October/November and Minimums in June/July. Precipitation patterns follow a very uneven spatial and temporal distribution, where most of the rains fall in the form of severe, very short and intense precipitation events during the wet season between December and March. These are particularly frequent between January and February, although seasonal rains are very erratic. The main inflows into the lake are the discharge from upstream and the local precipitation, estimated around 90% and 10% respectively (Kurugundla et al., 2018). Groundwater is disconnected from the upstream wetlands due to the geological faults separating the geohydrology of the two systems (Milzow et al., 2009; Monsimanyana et al., 2015). The main outgoing fluxes of the lake are seepage and direct evaporation (Mosinmayana et al., 2015), which is coupled with the inflow of water to the lake, during the warmest months. Therefore, water balance in the lake is mainly regulated by surface run-off from the wetlands and direct evaporation. This balance is slightly adjusted by inter-annual variations of local rains, and the local levels of groundwater. Flood failures upstream translate into sustained lake desiccation, whilst a continued series of wet years recharge groundwaters to sustain surface water over to the next flooding season (Hamandawana, 2008). The lake itself is an important indicator of the extent of the floods in the Okavango Delta.

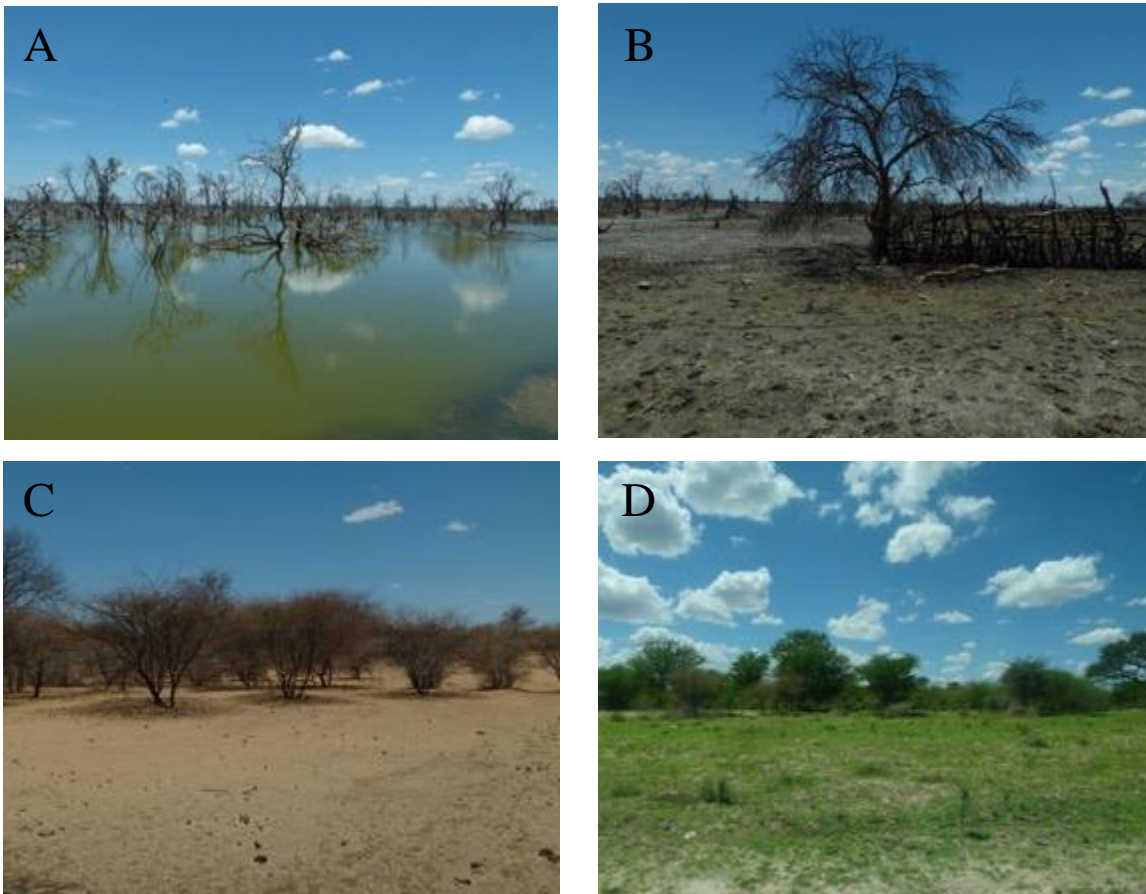


Figure 4: Photos of the vegetation in the floodplains of Lake Nagmi (A and B) during wet and dry periods, respectively. Vegetation conditions around Toteng during dry (C) and wet (D) season.

The study area includes the transition of different vegetation systems, from the open water in the lake to the floodplains of mopane trees (*Colophospermum mopane*), and to more mixed vegetation communities of thorn acacia woodlands and shrubs more inland from the lake. The lakebed area itself is a grassland during dry periods. However, and due to drought occurrence in the late XX century, shrubs and woodland vegetation encroached into the lake. *Cyperus papyrus* and other water plants have proliferated in the shores of open water during wet years. Surrounding vegetation is mainly represented by *Acacia eriloba*, typical in deep sandy soils of Southern Africa. Photos in figure 4 (A through D) show the landscape composition nearby Lake Ngami, where shrubs and sparse acacia trees dominate the vegetation cover. Patches of grasslands cover the ground during the rain period, being exposed to bare soil during the dry months of the hydrological year.

These surrounding landscapes of Lake Ngami sustain a local population of 7673 people mainly living in Sehithwa and Toteng (Botswana Central Statistics Office, 2011). The main economic activity is livestock rearing, a pastoralist activity strongly dependent on environmental conditions, however, in recent wet years fisheries and sporadic eco-tourism have gained importance. The lake is also part of the RAMSAR List of Wetlands of International Importance (UNESCO, 1971) and, when flooded, an important hub in migratory routes of wildlife in Southern Africa.

Table 1: Elevation profiles and total area of the buffer rings around the shore of Lake Ngami.

<b>Buffer</b>	<b>Area (km<sup>2</sup>)</b>	<b>Min (m)</b>	<b>Mean (m)</b>	<b>Max (m)</b>	<b>Stdv (m)</b>
1	100.7	922	929.6	952	3.3
2	94.8	923	934.7	962	8.2
3	99.1	923	938.9	968	10.9
4	105.8	925	940.7	972	11.9
5	111.9	925	941.7	971	11.9
6	118.1	924	942.1	974	11.7
7	124.3	925	942.5	975	12.1
8	130.6	925	943.1	977	12.6
9	136.9	925	943.6	983	13.1
10	143.1	925	943.9	987	13.2

Geographic buffer zones are useful to study distance effects. This analytical approach has been implemented previously for studies related to integrated water management and landscape ecology (Zhang et al., 2017). In order to assess the spatio-temporal effects of surface water variations on the dynamics of vegetation cover in the surrounding environments of Lake Ngami, it was applied equidistant zoning of 1 Km buffers from the maximum extent from the shoreline during the study period (Figure 3). This was decided to remove the effects of temporally flooded areas from the analysis of the vegetation cover derived from remote sensing data. The distance of 1 Km interval

ensures that at least 2 full-pixels were included perpendicularly in each buffer stripe. The interior-most ring covers an area of approximately 100 Km<sup>2</sup>, while the outer most buffer covers roughly 140 Km<sup>2</sup>. Table 1 shows the topographic characteristics of each of the buffers, including the areal extent. This way, a spatial discretization was created. The buffered zoning allows investigating the effect of a water availability gradient with relation to the vegetation growth, which is expected to be lower at the outer buffers compared to the inner shoreline of the lake.

## 2.2. Materials

Blending of remote sensing data sets for quantitative measurements of precipitation, water surface extent and vegetation is limited by the operational period of the data sets used. This is further aggravated by limitations in the time-series of in-situ hydrological data for discharge and precipitation. The identified period with the maximum available coverage for all the selected datasets was between 2000 and 2012. Availability of representative time-series for local data also played a role in the selection of the remote sensing products to maximize the overlapping of all data sets used. The study period was defined on the basis of all the time-series, as the maximum window of overlap between the entire collections of data sets selected (Table 2). This allows to verify and cross-check hydrological variables and assess the reliability of remote sensing products within the scope of this study. Also, the time-window from in the study-period represents a period including the transition from a minimum lake extent to a maximum recorded extent.

Table 2: Data availability, frequency and description of all datasets

<b>Data set</b>	<b>start</b>	<b>end</b>	<b>frequency</b>	<b>Description</b>
Discharge at Toteng	Jun 1, 1970	Jul 22, 2013	Manual	Water levels measured on E-type gauge plates during relatively regular fieldtrips by Department of Water Affairs staff.
Precipitation at Maun Airport	Oct 1, 1921	Mar 1, 2012	Monthly	Precipitation measured every 10 seconds, then averaged for 5 minute intervals using Vaisala integrated instrument. Station owned by Department of Meteorological Services.
MODIS EVI 16-Day Combined	Feb 18, 2000	Mar 14, 2017	Daily	This product is generated from the MODIS/MCD43A4 surface reflectance composites, using Near-IR, Red and Blue bands of each scene, and ranges in value from -1.0 to 1.0.
Global Surface Water	Mar 16, 1984	Oct 18, 2015	Monthly	This dataset contains maps of the location and temporal distribution of surface water from 1984 to 2015 on a month-by-month basis. The collection contains 380 images.
CHIRPS	Jan 1, 1981	Dec 31, 2018	Daily	CHIRPS incorporates 0.05° resolution satellite imagery with in-situ station data to create gridded rainfall time series.

### 2.2.1. Local data sets

Local data sets of meteorological and hydrological measurements were used as ground hydrological data. Both local data sets are obtained from a local repository managed by the Okavango Research Institute Monitoring Team (Okavango Research Institute, 2011). Discharge streamflow was recorded from the Kunyere river at Toteng, the only current discharge from the Okavango swamps into Lake Ngami (Figure 5). This data set measures discharge volumes in  $\text{m}^3/\text{s}$  and is recorded manually by members of the Department of Water Affairs with a monthly frequency during discharge periods. Although data is scanty for some months, it represents a source of ground information for the analysis.

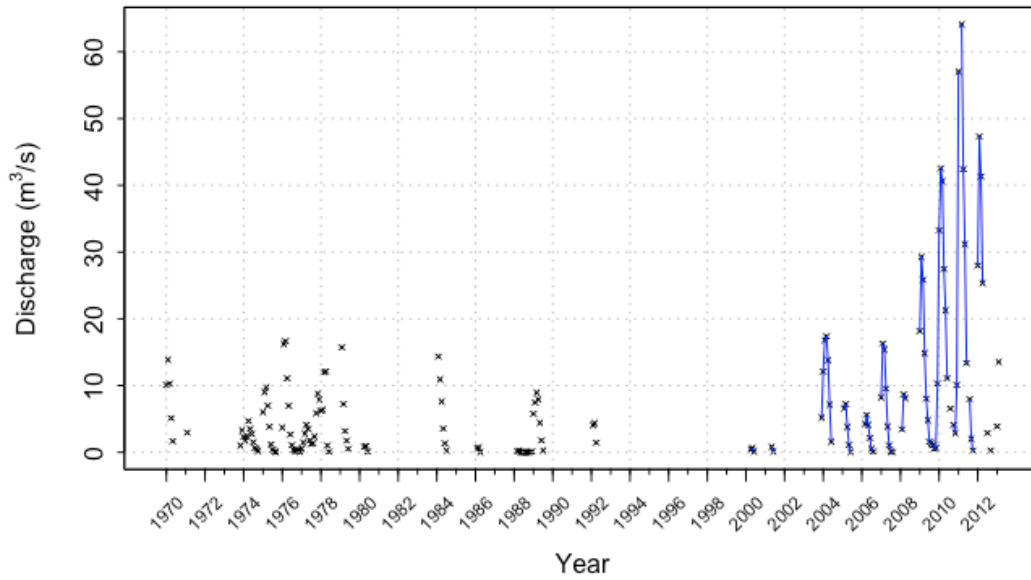


Figure 5: Entire time-series of the discharge ( $\text{m}^3/\text{s}$ ) at Kunyere river (inlet to Lake Ngami) from 1970 until 2013. The black crosses indicate discrete measurements of discharge. Blue lines represent continuous (monthly) measurements of the discharge during the study period.)

Local precipitation measurements are not available directly inside the defined study area. The closest meteorological station with consistent measurements of precipitation is located in Maun, approximately 80 Km from the lake. Figure 6A shows the full time-series of rainfall measurements at the meteorological station in the airport of Maun. The original daily measurements of precipitation were aggregated to generate a time-series of total monthly precipitation during the entire study-period (Figure 6B). Due to the characteristic erratic rainfall patterns in this region, the data cannot be assumed to perfectly describe the precipitation patterns in Lake Ngami. However, it serves as a ground reference for precipitation.

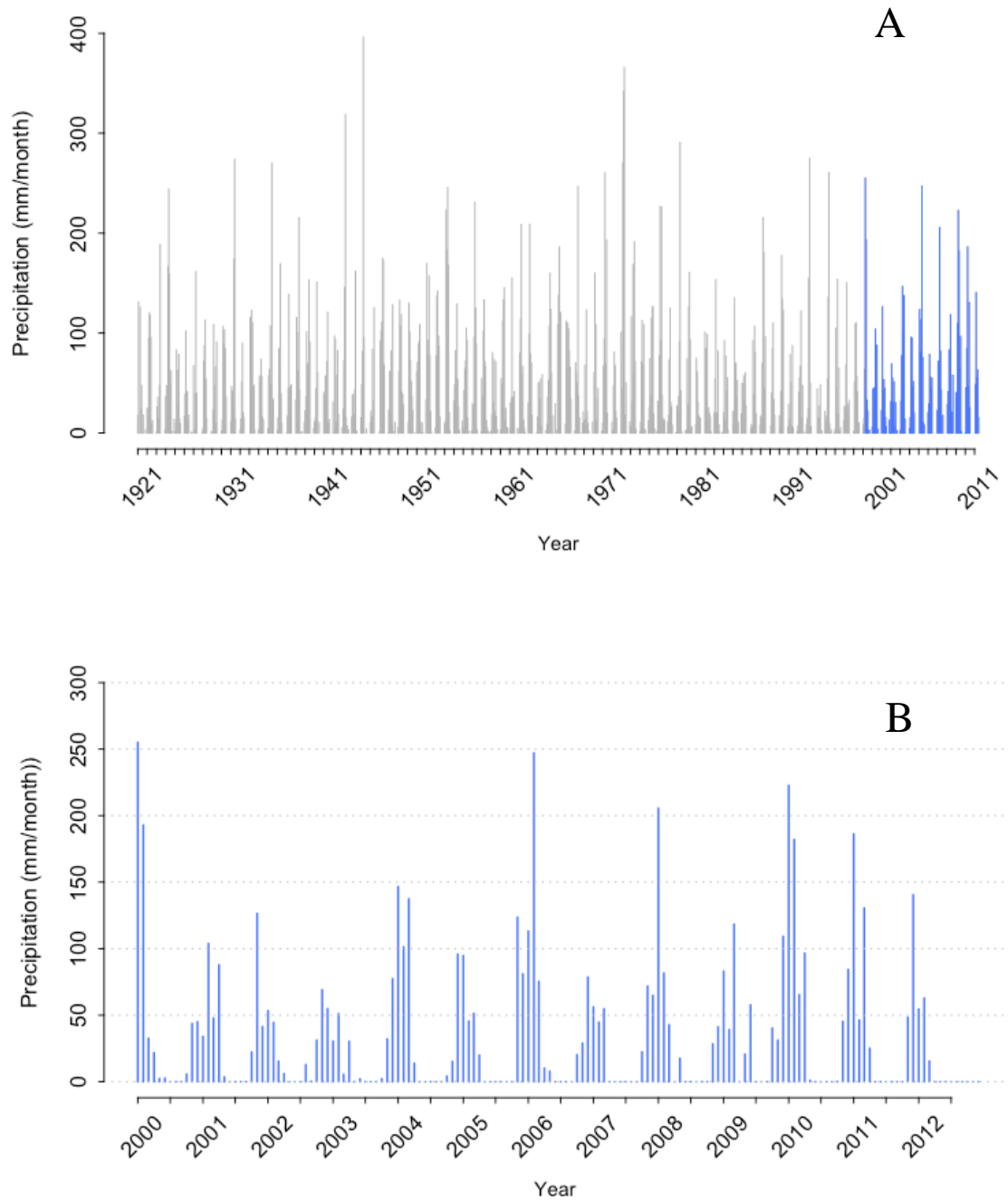


Figure 6: Precipitation time series at Maun airport. (A) Entire monthly precipitation time-series at Maun airport since 1921 until 2012. (B) Sub-set of the monthly total precipitation (mm/month) at Maun airport during the study period (2000-2012).



### 2.2.2. Remote Sensing of vegetation

Vegetation growth is a key component of ecosystem dynamics under changing environmental conditions. Vegetation responds to shifts in climatic conditions by adjusting physiological processes (e.g. photosynthesis) depending on water availability for the plant. One important characteristic of vegetation is its capacity to respond to seasonal variability, adjusting its growth to the seasonal conditions. This plasticity is a characteristic of the vegetation annual life cycle, generally referred to as phenology. Vegetation conditions are an important indicator in environmental monitoring due to its sensitivity to variations in environmental conditions. Moreover, plant phenology is strongly correlated with climatic drivers (Ma et al., 2013; Ge et al., 2016), determining the duration of the photosynthetically active period. In arid regions, plant phenology is also very responsive to water availability and hydrological fluctuations in the system which limit plant physiological activities (Jin et al., 2013).

Biophysical metrics of vegetation can be derived on the ground, at the plant level using field techniques. These methods provide phenological information of vegetation with a high temporal resolution, but are limited by spatial coverage. On the other hand, remote sensing data provides continuous measurements of reflectance from vegetation. Vegetation phenology result in seasonal variations of the energy reflected due to changes in the surface roughness and albedo of the canopy. This information can be monitored over the landscape or larger geographic areas, and sensors such as MODIS provide a relatively high temporal resolution with daily revisit times (Huete et al., 2002). Therefore, remote sensing data allows for quantitative measures of energy reflected, which can be related to biophysical features of vegetation. Healthy green vegetation has a characteristic reflectance curve defined by the absorption of red and near infrared (NIR) part of the spectrum resulting from the pigments and water in the canopy absorbing most of the energy around those wavelengths (Figure 7). In contrast, stressed vegetation has lower reflectance in the infrared due to lower water content in the plant tissues.

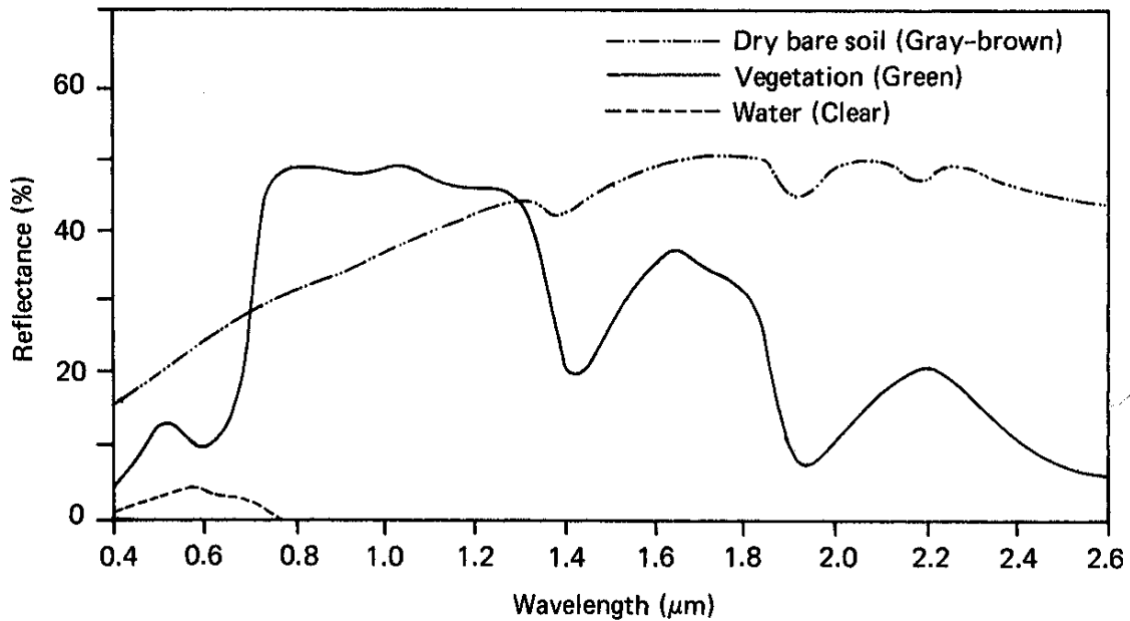


Figure 7: Typical spectral reflectance curves for vegetation, soil and water (Lillesand et al., 2014).

One of the primary methods in the use of Earth observation and satellite platforms is to monitor vegetation status using vegetation indices. These are arithmetical operations of pixel reflectance of two or more bands. Band ratios are used to monitor vegetation status and provide spatial and temporal comparisons of the status in the canopy, allowing to measure seasonal, inter-annual and long-term variations of vegetation health, productivity as well as phenological parameters (Huete et al., 2002, Ge et al., 2016). Vegetation indices derived from the reflectance of spectral bands have been widely used to monitor vegetation phenology. These indices aim at isolating photosynthetically active signals to differentiate vegetation activity across space and time (Huete et al., 2002).

For this study, the Moderate Resolution Imaging Spectroradiometer (MODIS) EVI 16-day composite products at 500 m were used from the Google Earth Engine data repository (<https://code.earthengine.google.com/datasets>). This is a “high-quality, cloud-free and filtered data” where pixels are selected on a basis of the best available pixel in the 16-day interval of the compositing period (Huete et al., 2002). The MODIS data sets were selected based on being temporally rich and having an acceptable 500 m pixel resolution for the study area in Lake Ngami (nearly 1500 Km<sup>2</sup>). This minimises the risk of confounding system variability with system change by reducing the sampling frequency and better characterizing vegetation phenology (Mishra et al.,

2012). Higher temporal resolution is also an important aspect since monitoring of vegetation phenology requires an appropriate frequency of measurements to capture temporal variations in phenology. Moreover, clouds and atmospheric conditions greatly reduce the usability of satellite imagery, limiting continuity of time-series derived data from platforms with lower revisit frequencies. MODIS provides mid-high spatial resolution with a daily revisiting time, and in this case, represents an adequate solution to detect vegetation temporal dynamics compared to other missions such as Landsat (Figure 8). On the other hand, vegetation in arid and semi-arid regions such as the study area, present a heterogeneous spatial distribution, generating mixed signals and limiting the sensor sensitivity to different land covers (Walker et al., 2014). Despite the general perception that higher resolution sensor (e.g.: Landsat) are more suitable for monitoring land cover (Munyati & Mboweni, 2013), there are complications related to the mixed signals from very fine resolution sensors due to increased signal-to-noise ratio in regions with scattered vegetation covers (Walker et al., 2014), complicating the retrieval of clearer dynamics due to higher noise (Figure 8B).

The Enhanced Vegetation Index (EVI) is a three-band index and uses the blue, red and near-infrared bands to assess various biophysical parameters and it has been shown to effectively track the seasonality in savannas (Jin et al., 2013; Ma et al., 2013). The values of this index range between -1.0 and 1.0, where a value above 0.1 generally indicated vegetation in the fallow season and values greater than 0.5 are related to a photosynthetically active period of vegetation (Huete et al., 2002). The EVI is calculated using the following formula:

$$EVI = G \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \times \rho_{red} - C_2 \times \rho_{blue} + L} \quad (1)$$

With the coefficients adopting the values:  $L = 1$ ,  $C_1 = 6$ ,  $C_2 = 7.5$  and  $G = 1$  (Huete et al., 1994). In this formula  $\rho$  accounts for atmospherically corrected reflectance,  $L$  is a coefficient to reduce radiance through the canopy, while  $C_1$  and  $C_2$  correct for aerosol influences in the red band. EVI also accounts for an atmospheric correction non-existing in other vegetation indices, resulting in a more stable measurement over long-term, by removing differences across seasons due to soil and atmospheric conditions. Temporal variations affecting atmospheric conditions such as sand storms (typical in arid regions) carry particles that interact with the sensitivity of the sensor. By correcting

for these conditions it is possible to remove atmospheric interactions from the data. The canopy correction coefficient ( $L$ ) is also relevant in open canopies where background signals reduce the canopy reflectance properties (Huete et al., 2002). At the same time, the Enhanced Vegetation Index follows a subtler response to photosynthetic activity, avoiding pixel saturation (a common feature in riparian floodplain vegetation), and improving pixel separation of different canopies (Murray-Hudson et al., 2015).

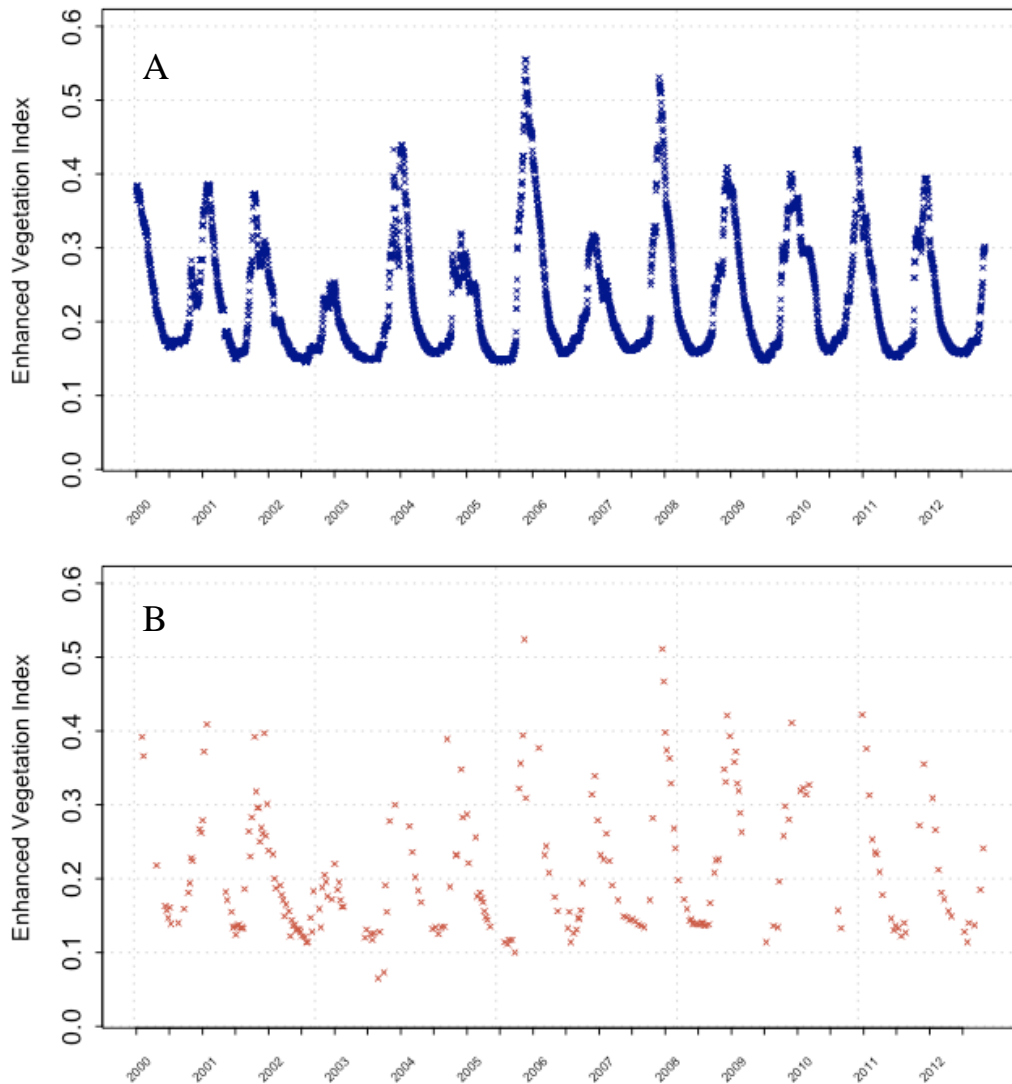


Figure 8: MODIS EVI 16-Day composite (A) and Landsat EVI 8 Day composite (B) mean values for the time series length between 2000 and 2013 in the Lake Ngami area.

### 2.2.3. Remote Sensing of surface water

Water availability is a determinant component of ecosystem functioning and structure, shaping landscapes and supporting vegetation growth. Spatio-temporal dynamics of surface water are even more relevant in drylands, where vegetation growth is supported by the fluctuations of this spatial and temporal variations in water presence (Rodriguez-Iturbe, 2000; von Hardenberg et al., 2001; Ma et al., 2013). However, accurate surface water detection is a complex task in multispectral remote sensing. Traditionally, spatial datasets for documenting open water location have relied on inventories and national surveys, statistical extrapolation of regional data and satellite imagery (Yamazaki et al., 2015). Variable conditions of water and its spectral properties in the surface of the Earth as a target of remote sensing are subject to a variety of conditions which affect the measurements of the sensor (Zhou et al., 2017). Figure 7 shows the typical spectral signature of water. However, this reflectance may be affected by plant growth, suspended sediments, and other water conditions, resulting in inaccuracies in the separation of water from other land surface features.

Global Surface Water data set (Pekel et al., 2016) represents a state-of-the-art global data set on surface water with 30 m pixel side resolution. It represents the first free and open-access global data set available of this type (<https://global-surface-water.appspot.com/>), allowing any user to obtain information of spatio-temporal presence of water at a global scale. The high accuracy of this data set (less than 1 % of false water detections and less than 5 % of omission) at a fine resolution, makes it a reliable source for monitoring surface water dynamics in open water bodies across the globe. This is a derived product from imagery including Landsat 5 Thematic Mapper (TM), the Landsat 7 Enhanced Thematic Mapper-plus (ETM+) and the Landsat 8 Operational Land Imager (OLI), spanning a period over 30 years between 1984 and 2015, with a measurement frequency of 16 days (Landsat revisit time). Several outputs are derived from this product, including the monthly recurrence of water detection showing the intra-annual distribution of surface water. The resulting data are binary grids of 30 m resolution, where each pixel is classified as open water when the totality of the surface is covered by water (900 m<sup>2</sup>), otherwise it is represented as land, or in some cases as a non-valid observation (null) for every image in the three Landsat missions.

The Global Surface Water data was complemented and cross-validated with the ground ancillary data of discharge into Lake Ngami. This approach allows contrasting the quality of a remotely sensed data in comparison to ground measurements. The uncertainty of the satellite data was also assessed (see Appendix II).

#### 2.2.4. Remote Sensing of precipitation

Measurements of precipitation beyond location-based observations are complex and involve several complications related to assumptions and interpolation artifacts. Gauging stations, despite the errors (sometimes around 20%), are the most direct measurement of rainfall. However, they lack continuous spatial information, which is provided by other methods such as radar or passive remote sensing. Usually, a combination of methods is the best approach to describe the spatio-temporal distribution of precipitation.

In Lake Ngami, precipitation is the other main influx contributing to the water balance in the lake due to the disconnection from the wetland's groundwater system (Milzow et al., 2009; Monsimanyana et al., 2015). Similarly to other environmental variables, the area lacks in-situ and readily available monitoring of precipitation. Although the police station of Sehithwa records precipitation, this data is not publicly accessible. The closest monitoring weather station on the ground with a relevant time frame covering the entire study period is located in Maun airport, around 80 km North-East.

Due to the lack of local rainfall data available, it is necessary to find other sources to estimate measurements of this variable in the study area. Satellite-based measurements such as the Tropical Rainfall Measurement Mission (TRMM) or the Climate Hazards InfraRed Precipitation with Stations (CHIRPS) provide means to obtain time-series of precipitation data where direct observations are not available (Huffman et al., 2001; Funk et al., 2015). Furthermore, remotely sensed data is spatially distributed, posing another benefit from this type of sources. CHIRPS provides a long-term record of precipitation with a daily frequency and a relatively high resolution (0.05°) (Funk et al., 2015), although this resolution of about 5 km pixel size at the latitude of the study area limits the discretization of precipitation amongst the different buffer stripes. It is therefore aggregated as the monthly sum for the entire area around Lake Ngami. CHIRPS data set integrates data from the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation

Analysis (TRMM) (Huffman et al., 2001) and combines it with measurements from gauging stations across the world. One of the public data streams feeding the CHIRPS gauge component is the Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL), which includes most of the weather stations from the Botswana Meteorological Service. Inclusion of weather station data from the region reduces the influence of interpolation artifacts in the study area of Lake Ngami. The units of the data set are in mm/day, later aggregated to monthly sums (mm/month) using reduction methods for the entire area surrounding the lake.

### 2.3. Methodology

The main goal of the present work is to test the viability of using the Google Earth Engine platform to generate consistent time-series, producing data sets to study the dynamics of vegetation and surface hydrology in the area of interest.

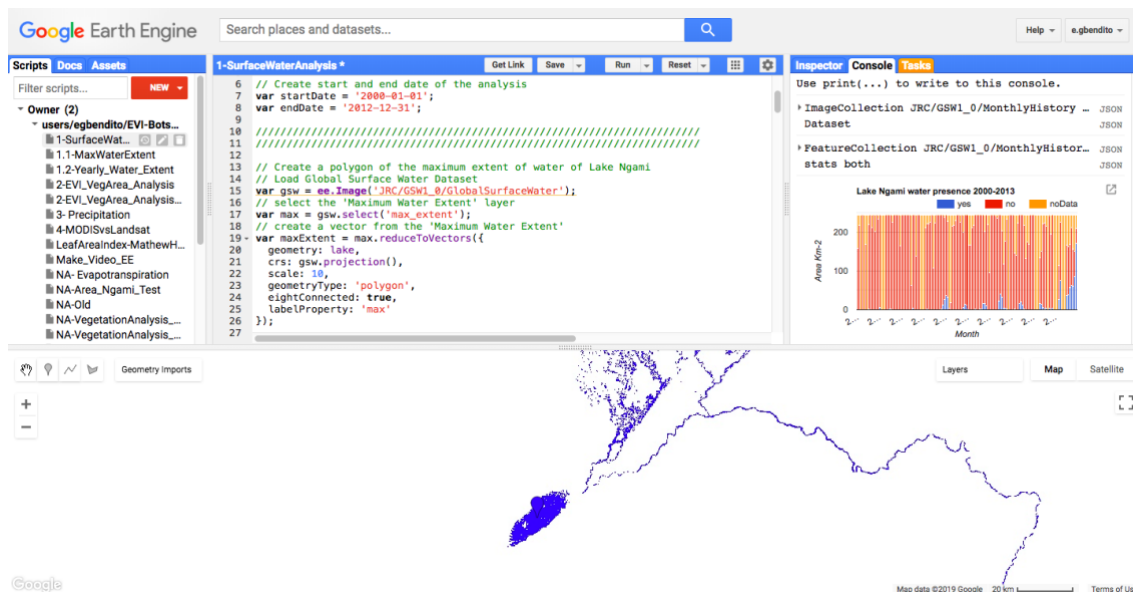


Figure 9: Web-based IDE of the application for the Google Earth Engine services, also referred as Earth Engine Code Editor (Google Earth Engine, 2019). The interface includes several features to develop geospatial workflows using JavaScript.

The above-described data sets (sections 2.3.2 to 2.3.4) are part of the repository included as a service in the Google Earth Engine platform. These are open-access and already available in the application. Furthermore, Google Earth Engine includes a set of libraries (or packages) to apply a variety of functions to handle the data. A caption of the integrated development environment (IDE) of the Google Earth Engine application is shown in figure 9. One of the most relevant techniques applied to the remote sensing data for this study are the so-called reducers. These are methods to apply a selected discretization on spatio-temporal data-series in order to create a regular and temporally representative metric of a variable over a time period and/or area of choice (Google Earth Engine, 2019). These functions allow to aggregate data over time and space from a full collection to a reduced version, comprising the dimensionality of the temporal domain (e.g.: monthly time series of precipitation from daily measurements). This method also extracts the comprised data-series over an area of interest, similarly to what is carried in other GIS software, and generally referred to as “zonal statistics”. This function aggregates data over a selected area, masking the values outside of it.

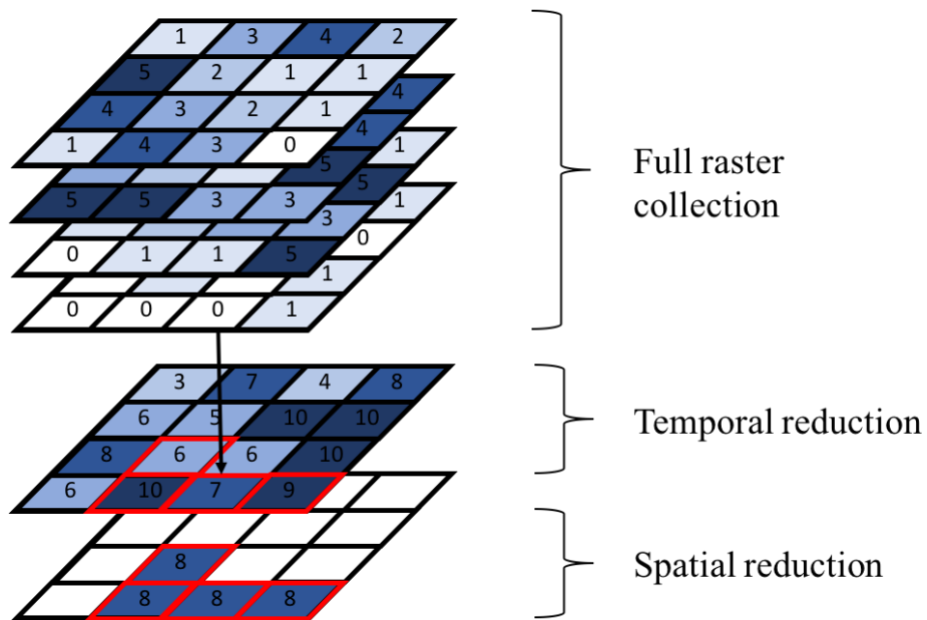


Figure 10: Schematic workflow representation of the reductions applied to the data carried on the Google Earth Engine platform. The original full collection of rasters were aggregated to monthly time-steps and later extracted to the area of interest.



Due to differences in the sampling frequency between the selected data sets (table 2), it was necessary to select a harmonized sampling frequency in order to create consistent time-series comparable across the different variables. For this reason, the temporal resolution for the study period was reduced to monthly aggregations. For instance, local precipitation data from Maun airport as well as the vegetation metrics from MODIS are provided with a daily frequency, while on the other hand, streamflow into Lake Ngami is recorded with a very inconsistent periodicity, depending on the seasonal discharge. To aggregate the different time-series in a comparable manner, the minimum temporal resolution possible was a monthly interval.

The combination of spatial and temporal reducers allows to systematically extract metrics for each of the 1 km buffers around Lake Ngami, with a monthly frequency spanning between 2000 and 2012. The scripts generated to obtain the reduced monthly data can be consulted in Appendix I. A schematic description of the workflow to achieve these aggregations is shown in figure 10. Typically, this approach returns a data set of reduced dimensionality for the different data sets. Water surface data is aggregated in the total area covered by water each month (in Km<sup>2</sup>). Rainfall is calculated as the sum of monthly precipitation in each pixel and then expressed as the average value of each buffer. EVI is aggregated as the monthly average value of all the pixels in each buffer. The final aggregated monthly data from these data sets over the entire study-period of 12 years provides 156 monthly measurements for each one of the variables.

Finally, the data was ported into the R statistical computing platform (R Core Team, 2017) in order to manipulate it and generate the statistics and graphs shown in the results section, using the packages ggplot2 (Wickham, 2016) and tidyr (Wickham & Henry, 2017). Figures 3 and 14 were created using QGIS version 3.2 (QGIS Development Team, 2016).



### 3. Results and discussion

#### 3.1. Local vs. Remote sensing data sets

Despite the mentioned increase in the use of remote sensing data to monitor and study ecosystems, it remains important to validate these data. Therefore, local data sets remain a key element to ensure that what is measured from space (or remotely) resembles the local situation. In the context of this study, local data sets were used to validate the satellite-derived dynamics of the two hydrological variables of interest.

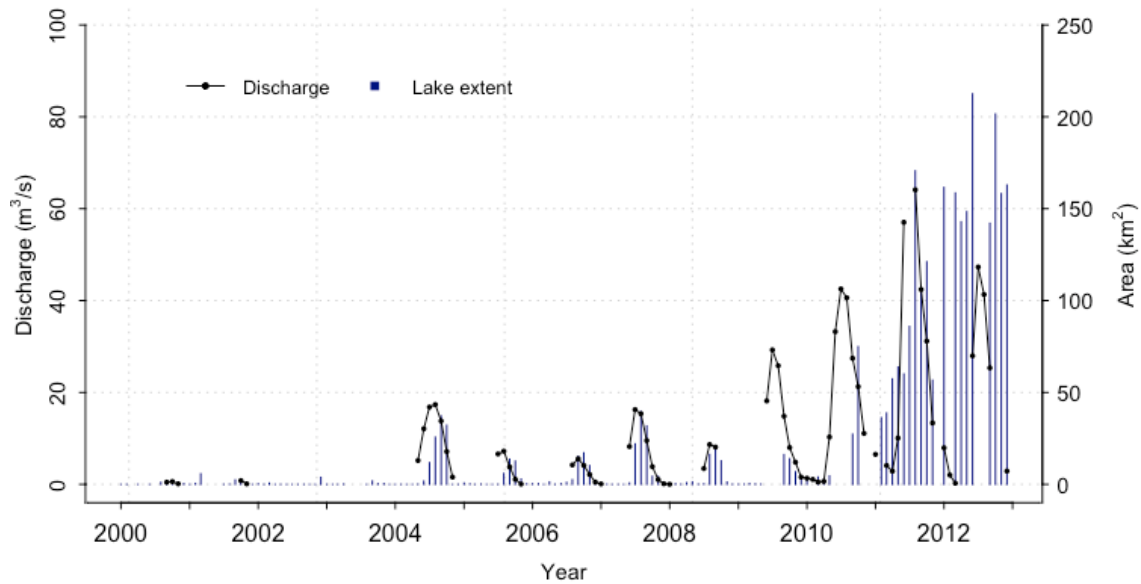


Figure 11: Time-series of the discharge (inflow) into Lake Ngami at Toteng showing also the barplot of the total extent of Lake Ngami on a monthly time-step between 2000 and 2012.

Discharge data is used as a proxy to describe the temporal dynamics of the flooding extent of Lake Ngami. At the same time, the Global Surface Water data set (Pekel et al., 2015) represents the actual spatial extent of the monthly coverage of open water in the lake. Figure 11 shows the monthly means for the run-off local dataset in comparison with those of the lake extent, based on monthly observations of water pixels. The extent of the lake matches with the inter-annual trend

of increasing discharge into the terminal Lake Ngami. Between 2000 and 2004 the lake was dry, disappearing completely in 2003 resulting in no measurements of discharge for that year, starting again in 2004 when the Kunyere and Nhabe rivers flowed into the lake. After 2010, the lake started to fill up from consecutive years of high discharge, reaching the maximum extent of 212 km<sup>2</sup> in June 2012.

Furthermore, it is important to highlight the coupled seasonality of the extent of flood pulses in relation to the incoming flows during sustained years of high discharge. As the flood wave arrives between May and June, the lake starts filling up. During the first floods, until 2008, the increase of the areal extent, is symmetrical to the inflows, with a month of difference between the highest inflow and the maximum extent of the lake for each year. Both the run-off and the lake extent decrease after this peak, which generally occurs around August each year, after which input from the river ceases, and the lake starts to dry, mainly due to direct evaporation (Mosinmayana et al., 2015). From 2009, an increase in the inflow results in the lake extent to increase remaining even during the months when discharge from the river stops. This is probably an indicator of the rise in the groundwater table, which in turn result in a larger extent volume of water remaining throughout the wet period between 2010 and 2012.

Table 3: Correlation matrix

	Dry period (2000 – 2008)	Wet period (2009-2012)	Entire period (2000 – 2012)
t	5.12	2.02	4.80
df	31	21	54
p-value	< 0.01	0.05	< 0.01
R	0.67	0.40	0.55

The correlation between the discharge data of the inflows and the water extent indicates that the area of open water in the lake is strongly determined by the inter-annual flooding events. Table 3 presents the correlation between the two variables. The correlation coefficient (0.67) and p-value ( $<0.01$ ) of the discharge and lake extent time series between 2000 and 2008 show a strong relationship. However, in consecutive wet years, as the lake fills up, the decrease in the run-off around August does not translate into a reduction of the lake area ( $p = 0.05$ ). During the wet years between 2010 and 2012, the correlation is lower, probably related to some gaps in the remote sensing data set due to clouds during the revisit date or to the recharge of the groundwater aquifer, sustaining water presence in the lake area over the year and replenishing the water loss to direct evaporation. Overall, the correlation between the local discharge data and the satellite data from the Global Surface Water data set indicates that remote sensing data may be used as a proxy to monitor water resources in a region such as Lake Ngami, which lacks a structured and continuous monitoring system.

Local precipitation and satellite-derived precipitation data sets were also confronted. Google Earth Engine provides the full data from the CHIRPS data (Funk et al., 2015), while the local measurements were derived from the repository at the Okavango Research Institute (Okavango Research Institute, 2018). The latter corresponds to local measurements at Maun airport, the closest available precipitation series. Both data sets were aggregated to represent the monthly sum of rain (mm/month), from which the yearly cumulative series were calculated (Figure 12).

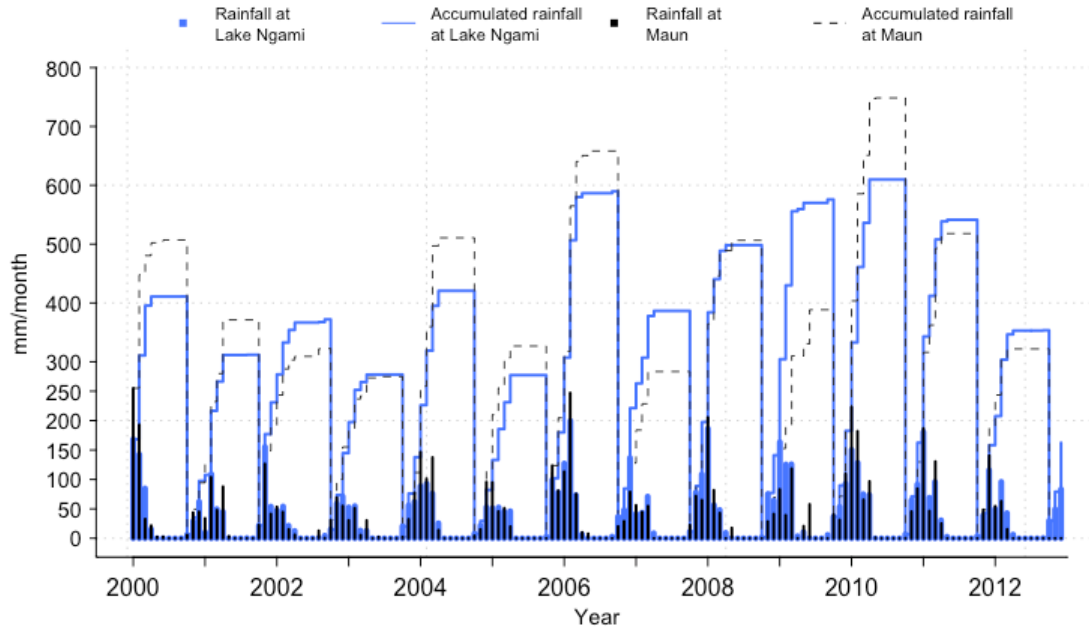


Figure 12: Comparison of the local data from a gauging station in Maun airport and the satellite derived precipitation at Lake Ngami. Both represent a monthly time-step including the cumulative precipitation between 2000 and 2012.

The seasonality of precipitation in Northern Botswana is strongly restricted to the rainy season starting between October and November and lasting until March of the following calendar year. Therefore, the hydrological year in the study area can be defined to start in October, based on the cumulative sums of precipitation. For the study period, the annual precipitation in Lake Ngami is 413 mm, compared to Maun of 410 mm. Monthly precipitation in Lake Ngami and Maun show similar patterns ( $R = 0.90$ ,  $p < 0.01$ ), although the monthly precipitation around Lake Ngami seems to be around 8% higher on average than the one in Maun during the 2000 – 2012 period.

The strong correlations and similarities in the rainfall patterns of both data sets suggest that CHIRPS may be a valid source of information to study precipitation in regions lacking a dense enough network of gauging stations. Moreover, the fact that CHIRPS includes ingested data from weather stations for calibration of the satellite-derived data, makes it a more integrated data set, accounting for the spatial component derived from satellite systems and the site precision of gauging stations.

### 3.2. Surface hydrology dynamics

The two hydrological inputs of Lake Ngami (precipitation and run-off) are derived using the remote sensing data from the Google Earth Engine repository, namely the Global Surface Water data set (Pekel et al., 2015) and CHIRPS (Funk et al., 2015). As stated above, these two data sets resemble accurately the hydrological conditions in the study area. Figure 13 presents both the dynamics of the precipitation and flooding in and around Lake Ngami between 2000 and 2012. As is generally stated in other studies (Milzow et al., 2009; Monsimanyana et al., 2015; Kurugundla et al., 2018), Lake Ngami flooding is virtually independent of local rainfall, and the filling up of the lake mainly depends on the run-off conditions from the upstream wetlands.

Overall, the precipitation around Lake Ngami during the study period does not seem to present significant differences from the long term annual mean rainfall in Maun ( $p > 0.05$ ). Variations in the amount of precipitation in the study area seem not to have an influence in the increase of surface water in the lake. Rather, discharge from upstream is the key variable that regulates the flooding conditions in the lake, as stated in other studies (Monsimanyana et al., 2015).

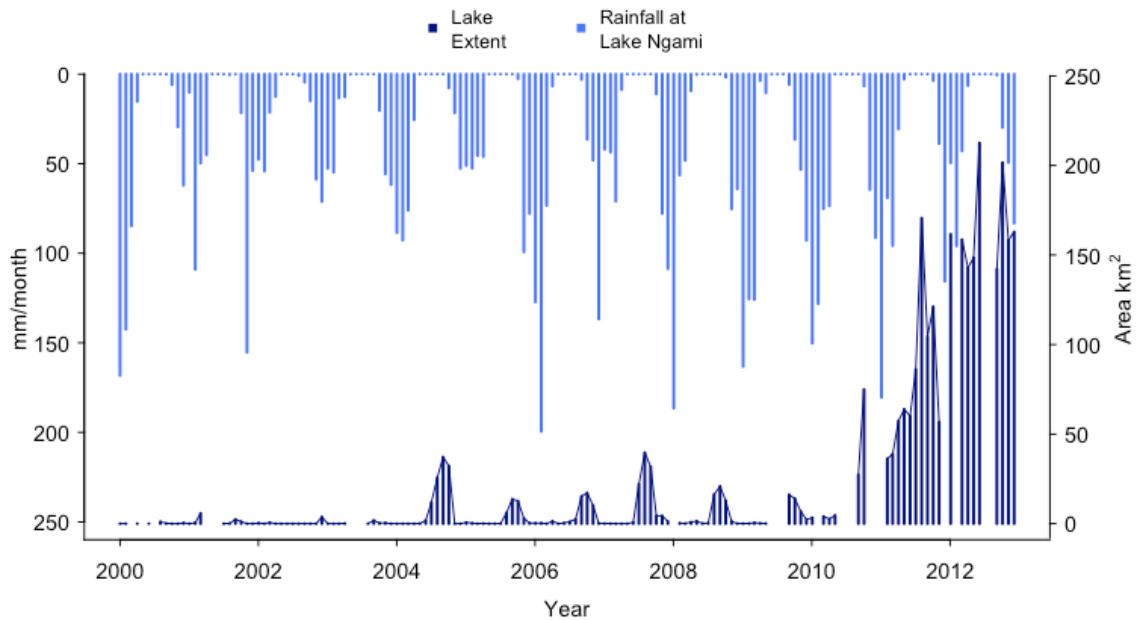


Figure 13: Monthly observations of lake extent (in  $\text{km}^2$ ) and precipitation (mm/month) in the Lake Ngami between 2000 and 2012.

Regarding seasonality of the hydrological processes, the results indicate that the flooding of the Lake Ngami is decoupled from local precipitation in the area, with opposing intra-annual dynamics for both variables. This is related with the seasonal precipitation in this part of the African continent, where precipitation follows a modal distribution between October and April (Gaughan et al., 2012; Kling et al., 2014). As a result, precipitation in the Okavango region does not support the flooding extent of the lake. Instead is the run-off coming from the upstream Angolan highlands which regulates the flood pulses in Northern Botswana, flowing for almost 6 months before reaching the Southern end of the Okavango basin in Lake Ngami (Milzow et al., 2009). In more detail, the flood wave into Lake Ngami arrives between May and June each year, followed by an increase of the flood pulse through the following four months, filling up the lake bed during the dry winter season in Northern Botswana. The inflow to the lake reaches its peak in August, followed by the maximum annual extent of the Lake around September/October. At this time, the hydrological year begins with the first rains of the wet season. This is the typical hydrological behavior of Lake Ngami, as described by other studies (Gondo and Pedzisai, 2016; Kurugundla et al., 2018). The results indicate that the Lake Ngami hydrological system is independent of the local rainfalls, as indicated by high rainfall years in 2006 with no effect on the extent of open water in the lake. However, 2008, 2009, and 2010 were above average wet years, after which the lake extent increased, probably related to a parallel high rainfall in the Angolan highlands during those years (Kling et al., 2014).



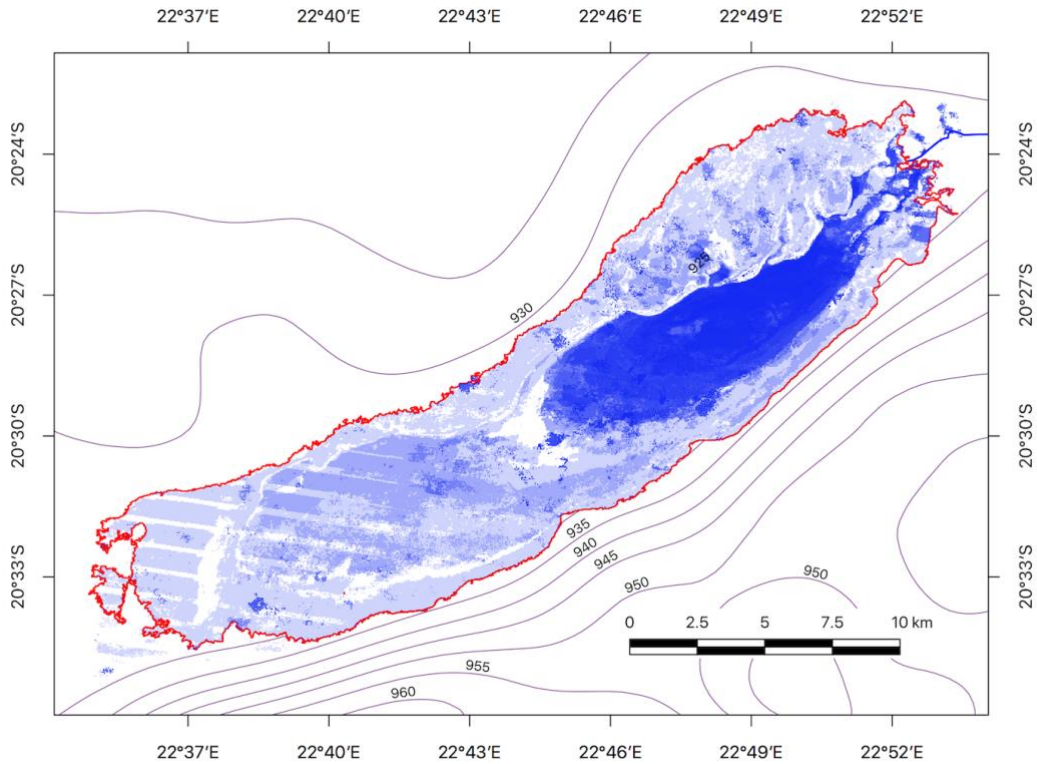


Figure 14: Annual transitions of open water in Lake Ngami between 2000 and 2012. Lighter blues represent lower frequency of water coverage. The red solid line indicates the maximum extent of the lake.

Spatial distribution of water presence based on monthly temporal profiles over an entire year allows to reconstruct the water extent dynamics of Lake Ngami as a union of the monthly measurements for a given year. Between 2000 and 2012, during the study period, there is a high variability of spatio-temporal distributions of surface water in the lake area. In general, the temporal variations in the lake are related to two main sub-periods: a relatively dry series of years between 2000 and 2009 and a high flooding period from 2010 until the end of 2012. The spatial and temporal distribution of the water in the bed lake during the study period is shown in figure 14. The annual extent of water in the lake between 2004 and 2009 is usually concentrated in the North-East section, corresponding to the typical coverage of open water reported by historical literature (Meier et al., 2015). The open water of the lake usually spreads over an area from 10 to 50 km<sup>2</sup> between those years (highlighted in intense blue in figure 14). During the wet period, corresponding to the years 2010, 2011 and 2012 the extent of the lake increases. In 2012 the maximum area covered by water is reached during the study period, spanning over 212 km<sup>2</sup> and probably representing the largest extent of the entire historical record of Lake Ngami. White areas

in figure 14 are likely related to the topography of the lake bed, resulting in very shallow waters (less than 0.5 m) and/or concentrations of drowned vegetation. However, some of those areas result from image scanning errors of the Landsat sensors (stripes in the image). The flood extent itself follows the elevation contours closely, and is possible that small tectonic variations change the morphology of the lake tilting the water inside the lake to different areas. This is a very recurrent explanation for the overall configuration of the Okavango basin (Wolski & Murray-Hudson, 2006; Milzow et al., 2009).

### 3.3. Vegetation dynamics

Inter-annual vegetation dynamics are shown in figure 15A for all the buffer zones. These series are the daily values of the 16-day composite of EVI from MODIS over the entire study period. All the zones present similar patterns, with slightly different EVI values across the different buffers. In 2006 and 2008 the vegetation reached the highest EVI peak values of the entire series, whereas 2003 is the year when the vegetation presented a very limited growth. Variability in the inter-annual development of vegetation is a typical characteristic in this type of semi-arid and arid climates, where the growth of vegetation is determined by water availability (Jin et al., 2013; Ma et al., 2013). Similarly, the several peaks within some of the years are probably a response to explicit rainfall events later in the rainy season. Overall, the results seem to indicate the high variability of vegetation growth in the Lake Ngami area from one year to another.

To further investigate the vegetation, average intra-annual dynamics were calculated. For each buffer zone in the study area, monthly EVI profiles were reconstructed using the mean value over the entire 1 km stripe and study period. This approach reconstructs the long-term intra-annual vegetation dynamics for each one of the 10 buffers around Lake Ngami (Figure 15B). These values are the daily aggregated means of each buffer for the entire study period, and therefore, represent the average vegetation response between 2000 and 2012 for each day of the year. Each of the buffer stripes varies in EVI amplitude values with minimums around 0.16 and maximums around 0.37. Generally, the further away from the lake, the lower the magnitude of change between the minimum and maximum EVI value.

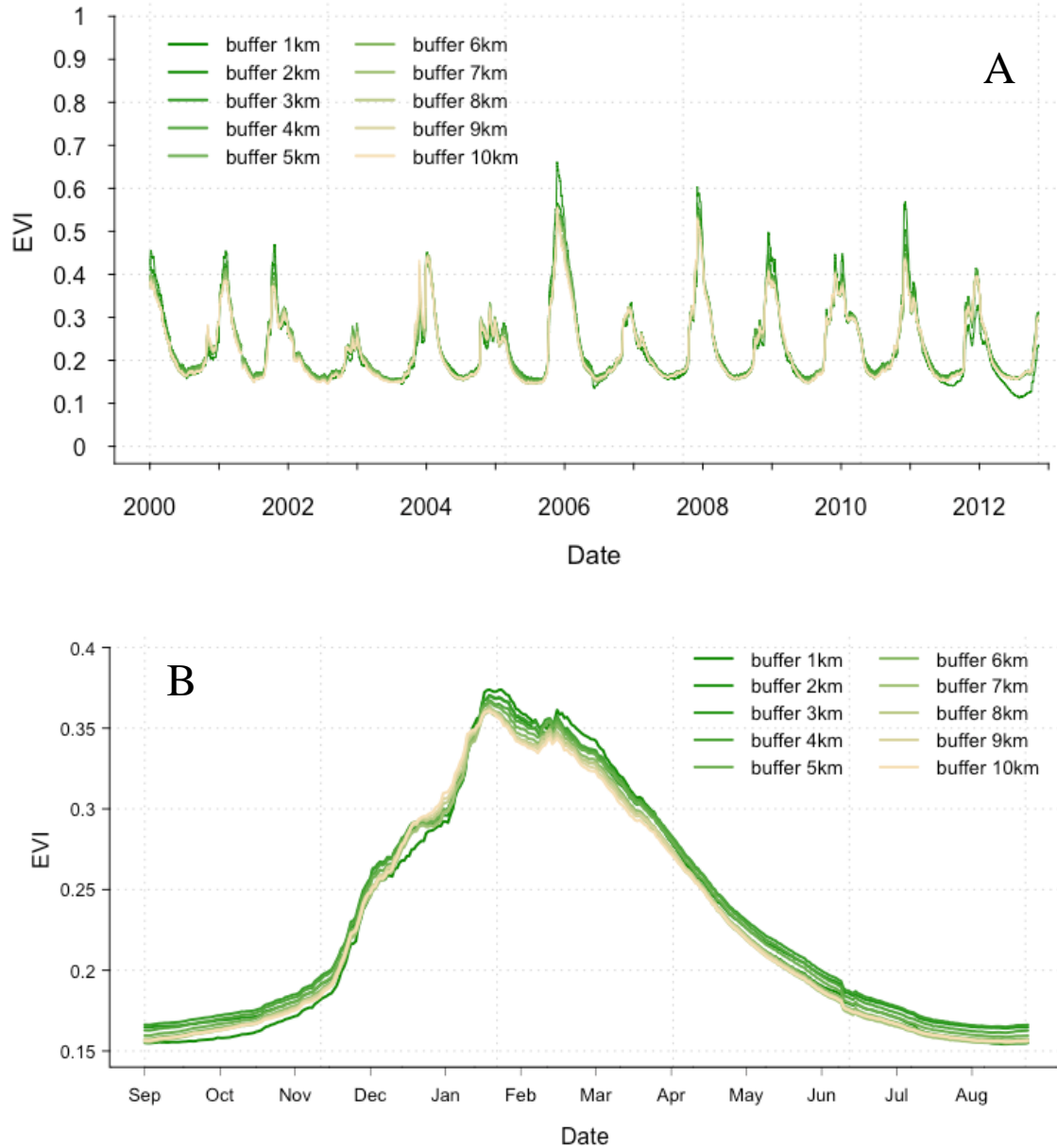


Figure 15: Time-series profile of the Enhanced Vegetation Index (EVI) during the study period from 2000 to 2012. (A) Each profile corresponds to a 1 km buffer stripe around the shore of Lake Ngami. Greener colours represent a buffer stripe closer to the shore of the lake, and the yellow ones the edge of the study area. (B) Average phenological seasonality of the vegetation dynamics for each buffer zone during the entire study period (2000 – 2012). The different lines correspond to the annual time series of each one of the buffers surrounding Lake Ngami.

The long-term average phenology of the vegetation in the area (Figure 15B) is similar for most of the contiguous buffer strips, reaching its peak in late January, with a second spike of vegetation growth around a month later. After this, vegetation begins the senescence stage, reaching the minimum EVI in late August. The Southern hemisphere winter months (June - August) show the minimum EVI values, while the period between January and March is the time with the highest levels of photosynthetic activity for vegetation in the study area.

In this area, the landscape is very patchy, being a mix of woodland savannas with shrubs and areas of open grasslands during the rainy season. As a result, EVI values are a general representation of the vegetation conditions over the 250m pixel width of the MODIS spatial resolution. However, this is accounted in the formulation of the EVI through the canopy coefficient ( $L$  in formula 1). Regarding the differences in phenology with increasing distance from the lake shore, it is possible to notice the effect of the distance from the lake in figure 15 (A and B), indicating some effect of the water availability in relation to the distance of vegetation from the lake. Vegetation closer to the lakeshore shows persistent higher EVI values than the more inland vegetation due to the higher availability of water. This is the case both in the inter and intra-annual vegetation patterns. In this sense, an interesting case is the immediate area of the lake shore, where the EVI curve results in the lowest values of EVI in 2011 and 2012, coinciding with Lake Ngami expansion. A similar result was found by Zhang et al. (2017) in semi-arid watershed in Northern China, where vegetation in floodplains showed lower photosynthetic activity. This could be related to two main reasons:

- a) During desiccation periods, shrubs and acacias encroached the previously grass-dominated areas. As the water levels increased during the study period, many of these plants probably were affected due to the saturation of the lower horizons of the substrate (Meier et al., 2015).
- b) Mixed signal of water and vegetation in the pixel, resulting in lower EVI values because of low reflectance in the infrared and near-infrared part of the spectrum (Zhou et al., 2017).

### 3.4. Coupling of hydrological and vegetation dynamics

To examine the effects of the variation and dynamics of hydrology on the vegetation around Lake Ngami, the two key phenological dates of minimum EVI and maximum EVI were identified for each year within the study period. Figure 16 presents the greening season of vegetation with the peak and minimum of the phenology for the year 2006 – 2007 (other individual years are shown in Appendix III). Typically, the greening season in the area runs between August and February (respectively lowest and highest annual EVI values), with small shifts depending on the start of the hydrological year. Moreover, minimum and maximum EVI metrics indicate the month in each year when the vegetation transits from the annual senescence, a period of decreasing photosynthetic activity, to the greening phase, the period of increasing photosynthetic activity. These are two relevant dates in the annual phenological pulse of vegetation in the Okavango region, as they reveal the increase of transpiration, with higher ratios of evaporation to transpiration when vegetation is less active (Dincer et al., 1987).

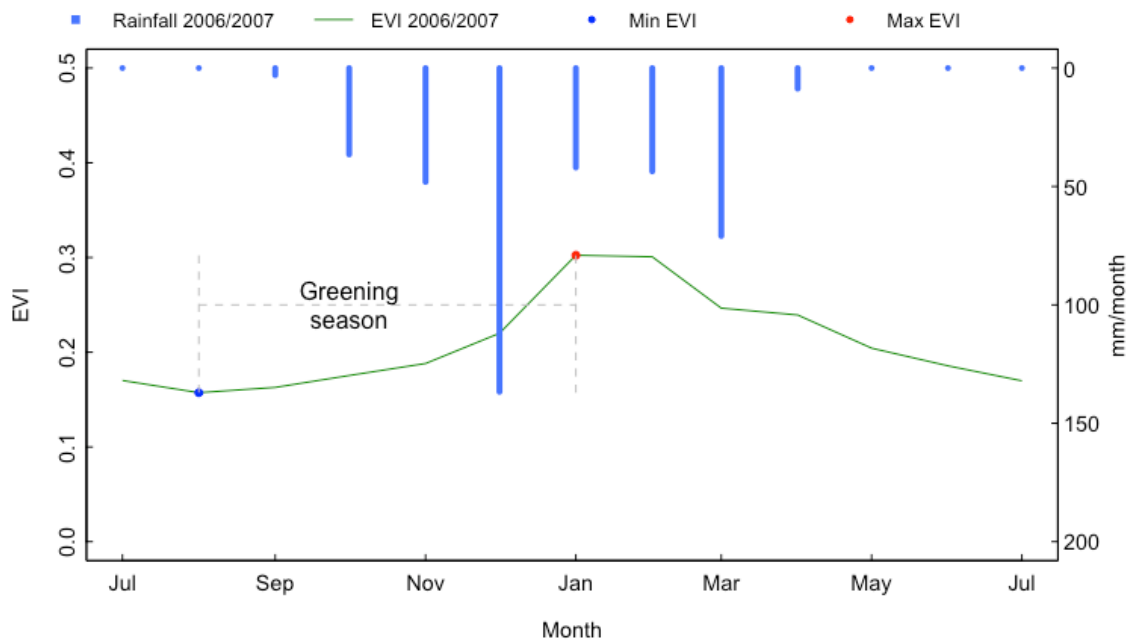


Figure 16: Monthly mean EVI composite over the whole study area around Lake Nagmi for the growth cycle of 2006 – 2007 with the valley and peak of the growing season (minimum and maximum EVI respectively). Monthly precipitation is included in the plot to depict the length of the hydrological year and volume of rain that season.

Figure 17, is a zoom-out of figure 16, and represents the full time-series of vegetation and rainfall dynamics with the average mean of the EVI reconstructed as a composite from all the buffers around Lake Ngami. It is reported together with the monthly amount of precipitation during the whole study period. Monthly vegetation dynamics around the lake follow an increasing trend within the hydrological year, matching the arrival of the first rains in November. In most years, the peak of the photosynthetic activity of vegetation is in late January to early February, and 8 out of 12 years comes with a month of lag after the wettest month of the year, generally January. After the maximum annual EVI is reached, greenness decreases, and the vegetation fallows. This seems to indicate that vegetation in this area is well adapted to the local precipitation patterns, and uses the beginning of the rainy season to set on the annual greening, perfectly matching the highest precipitation month to reach the maximum of its annual development. This is a typical characteristic of savanna vegetation in Northern Botswana (Jin et al.,2013) and other ecosystems under arid and semi-arid climates (Ma et al., 2013), indicating the adaptation of vegetation in arid and semi-arid regions to the dynamics of seasonal rainfall, as well as its importance in ecosystems under those conditions.

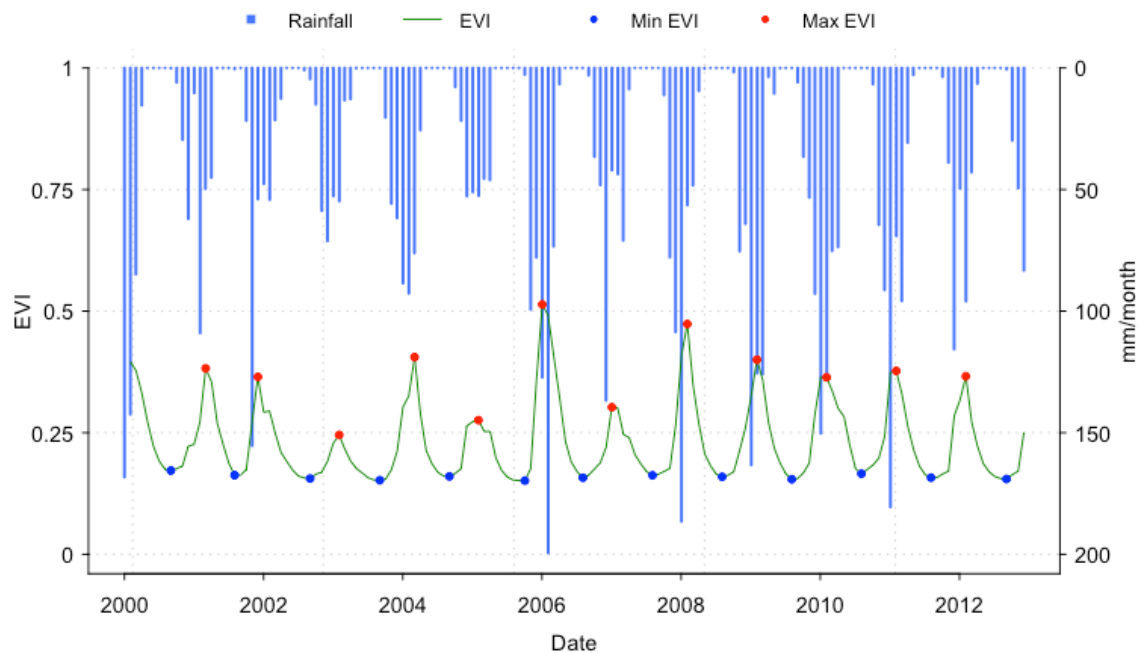


Figure 17: Time-series of the monthly precipitation and vegetation EVI from 2000 to 2012 around Lake Ngami. The EVI profile corresponds to the aggregated means of all the buffers composing the study area. This includes the annual minimum and maximum EVI values.

In order to study the effects of annual precipitation in relation to vegetation over each entire hydrological year, annual maximums of EVI together with accumulated precipitation are presented in figure 18, indicating some correlation ( $R = 0.64$ ,  $p < 0.05$ ). During the study period, maximum EVI values fluctuated between 0.24 (2003 - 2004) and 0.51 (2006 - 2007), with a total accumulated rainfall of 275 and 658 mm, respectively for each year. This suggests that maximum vegetation growth in a year may be driven by the annual incoming rainfall, as is the case in other dry savanna regions of the world (Ma et al., 2013).

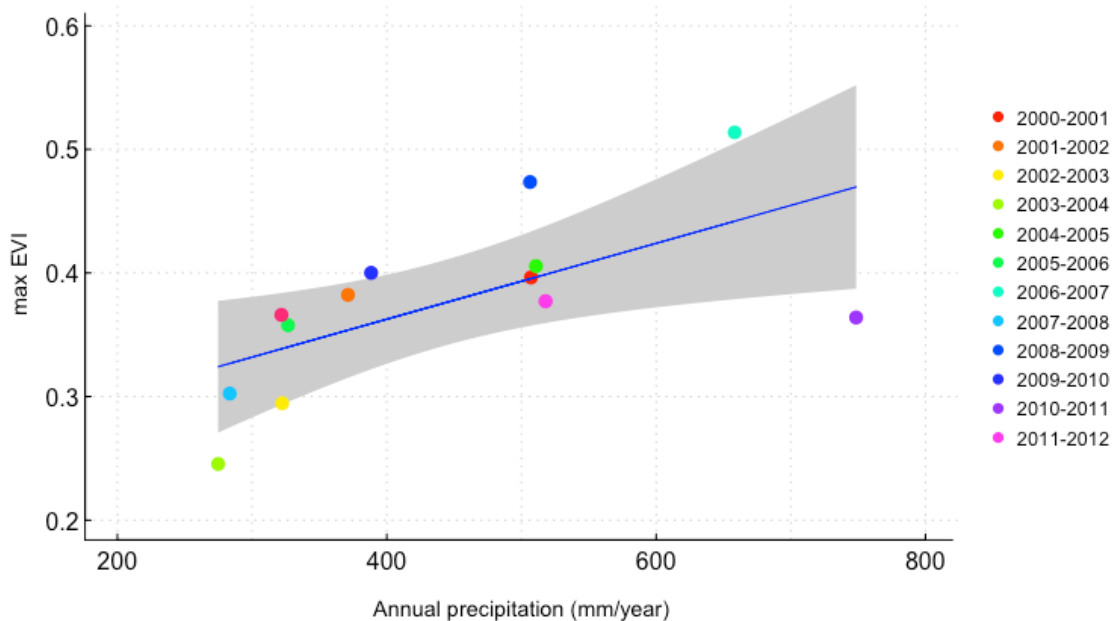


Figure 18: Relationship of each annual precipitation (mm/year) and maximum EVI derived from the average phenology in the surrounding area of Lake Ngami during the 2000 - 2012 study period. The shaded area indicates a confidence interval level of 95% and the blue line is the fitted linear model of the two variables.

To investigate the response of vegetation to surface hydrological inputs (lake extent and precipitation) with increasing distance from the lake, the monthly mean EVI for each one of the increasing distance buffers around the lake, was compared separately with the two hydrological variables of interest. Months occurring during the greening season of each year were selected for this purpose, creating a subset of 72 measurements of EVI, precipitation and lake area for each of the 10 buffer stripes. This data represents nearly half of the full monthly time-series. The selection of the greening phase was based on the assumption that during those months the water requirements are higher, and therefore the vegetation will benefit from the water available from the different sources (lake and precipitation).

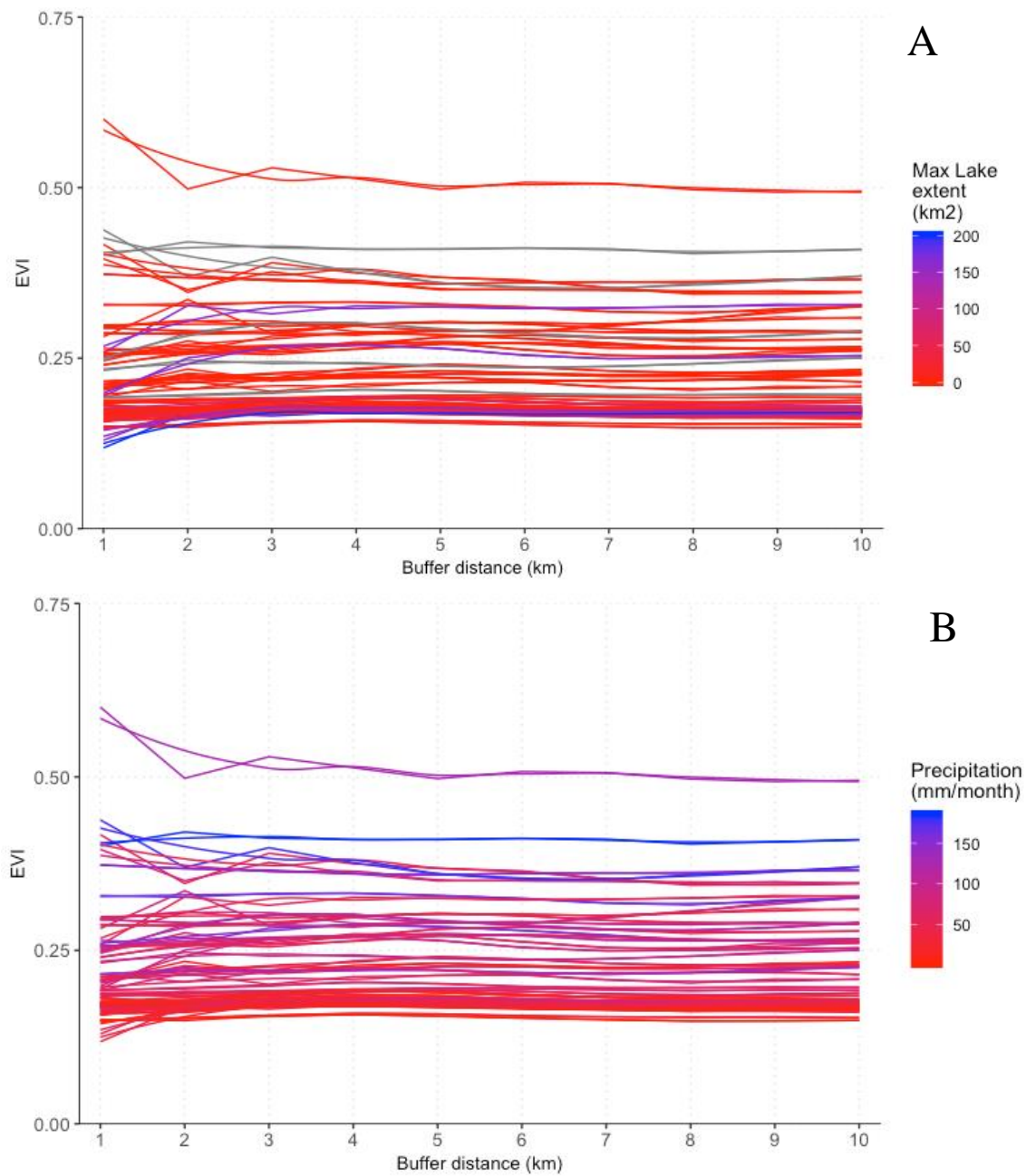


Figure 19: Comparison of the response of vegetation (EVI) to the lake extent (A) and monthly precipitation (B) at each of the contiguous buffers around Lake Ngami. The lines represent each month of the greening season between 2000 and 2012.



Figure 19 shows the mean EVI in buffer zones of different distance to the lake, and the relationship of vegetation with the monthly conditions of the lake (A) and precipitation (B). Differences in the response of vegetation to each variable are clear. Precipitation around Lake Ngami has a greater impact on the growth of vegetation, resulting in larger EVI values across the 10 km distance as the monthly precipitation increases. Meanwhile, vegetation does not seem to respond to the effects of vicinity to the surface water of the lake by increasing its growth, and larger lake areas result in indistinctive response from the vegetation compared to dry periods when the lake was dry.

Across the distance of the 10 buffers, vegetation response to hydrological factors is variable. Several months present the situation where, within the first kilometer near the shore of the lake, EVI is low, after which it increases. In most cases, beyond 5 kilometers from the lake EVI stabilizes. These results are similar to the study by Zhang et al. (2017), who investigated the riparian vegetation in relation to flood events in a watershed in North-West China. In Lake Ngami however, some cases indicate that closer to the lake EVI values are higher. These are usually related to extreme precipitation months, generally over 100 mm. Notably, the maximum extent of the lake result in lower EVI values. This could be related to the effects of soil saturation or to the mixing of water and vegetation signals as mentioned previously. This situation can also be observed on figure 15 A, where during 2012 the EVI within the first kilometer from the shore drops to the minimum of the entire study period, coinciding with the maximum extent of the lake.



## 4. Conclusions

The present study has carried out the investigation of the two main objectives framed in section 1.5. As a main conclusion, it highlights the usefulness of a platform such as Google Earth Engine for storage and preprocessing of remote sensing data to monitor hydrological and vegetation dynamics in areas with scarce local monitoring networks, such as is the case of Lake Ngami.

Currently, remotely sensed data represent one of the main sources of information on environmental variables. It is to a large degree free accessible, and new platforms such as the Google Earth Engine provide ready and preprocessed data sets for a wide range of applications. This platform in particular represents an important resource to deal with this kind of data sets. Satellite data is currently considered as a form of big-data, and requires of costly storage spaces. Furthermore, processing and analysis of this data is very time-consuming to correct and rectify the original data under average computational power. Google Earth Engine provides already pre-processed data, along with a powerful infrastructure and extensive library to further interact with the data and extract it for further analysis. This represents a tremendous advantage to study bio-geophysical processes in ecosystems without the associated costs of handling this type of data.

Moreover, the data sets used in this study provide relevant information in areas lacking the necessary monitoring networks, and correlated well with local measurements of hydrological variables. This is an important added value from these products, as it allows to reduce the costs of extensive monitoring networks, and provides long-term monitoring of environmental variables with a high level of accuracy, as was shown through the validation of satellite products with the local precipitation and discharge data.

The above-mentioned materials allowed to investigate the vegetation and surface hydrology dynamics around Lake Ngami. First, the hydrological settings of Lake Ngami result in very dynamic flooding patterns with large variations in the extent of the lake between 2000 and 2012. The system relies entirely on the inflows from upstream, and is virtually independent from local rainfall to sustain water levels in the sub-basin. On the other hand, vegetation is mainly dependent on the patterns and seasonality of the precipitation in the area. The annual growth of vegetation responds readily with the arrival of the first rains, setting off the beginning of the growing season,

with an increment in the photosynthetic activity of vegetation. The maximum growth of the season usually corresponds to a high rainfall month, indicating that vegetation seems to be well synchronized with seasonal rainfall patterns, boosting its growth during the wettest period of the year. Regarding the relationships between flood extent and vegetation growth, evidence was mixed. Initial results indicate that vegetation near the lake yields higher EVI values. However, further investigation reveals that the annual flooding of the lake during high discharge years does not seem to support vegetation in the immediate surroundings. Even in high flood years, vegetation followed similar dynamics, reaching minimum EVIs similar to low flood years. It is more likely, however, that the flooding of the lake is more related to groundwater dynamics, which could sustain vegetation in prolonged periods of drought.

In general, it is important to consider the effects of climate change scenarios and variations in the temporal rainfall patterns in this region. Increased lake extent has more relevance for the local population, as it provides recharge to the local boreholes, acting as a natural reservoir for low rainfall years. Although quite dynamic and resilient, Lake Ngami is a very fragile ecosystem, susceptible to extensive impacts deriving from climate change and more erratic rainfall patterns upstream, as well as to human interventions and flow diversions such as hydropower or irrigation systems. Albeit vegetation dynamics in the area show some plasticity, phenological patterns are likely to be affected under sustained dry periods of low rainfall, impacting the livelihoods of local communities.

## 5. References

- Akanyang, P. (1997). The geology of Lake Ngami area: An explanation of Quarter Degree Sheet 2022B. Lobatse, Botswana: Geological Survey Department, Ministry of Mineral Resources and Water Affairs.
- Botswana Central Statistics Office (2011). Population and Housing Census, 2011. Gaborone: Statistics Botswana. Retrieved from [https://www.webcitation.org/68QAYaEFm?url=http://www.cso.gov.bw/templates/cso/file/File/Population%20of%20Towns%20and%20Villages%20\\_2011%20PHC\\_%20Jun%204%202012..pdf](https://www.webcitation.org/68QAYaEFm?url=http://www.cso.gov.bw/templates/cso/file/File/Population%20of%20Towns%20and%20Villages%20_2011%20PHC_%20Jun%204%202012..pdf)
- Choler P., Sea W., Leuning R.; 2011. A Benchmark Test for Ecohydrological Models of Interannual Variability of NDVI in Semi-arid Tropical Grasslands. *Ecosystems*, 14, 183-197.
- Dincer, T., Child, S., & Khupe, B. (1987). A simple mathematical model of a complex hydrologic system—Okavango Swamp, Botswana. *Journal of Hydrology*, 93(1-2), 41-65.
- D'Odorico, P., Laio, F., Porporato, A., Ridolfi, L., Rinaldo, A., & Rodriguez-Iturbe, I. (2010). Ecohydrology of terrestrial ecosystems. *BioScience*, 60(11), 898-907.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data*, 2, 150066.
- Gaughan, A. E., Stevens, F. R., Gibbes, C., Southworth, J., & Binford, M. W. (2012). Linking vegetation response to seasonal precipitation in the Okavango–Kwando–Zambezi catchment of southern Africa. *International journal of remote sensing*, 33(21), 6783-6804.
- Ge Q., Dai J., Cui H., & Wang H. (2016). Spatiotemporal variability in start and end of growing season in China related to climate variability. *Remote Sensing*, 8(5), 433.
- Google Earth Engine. 2019. Retrieved from: <https://developers.google.com/earth-engine/>

- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18-27.
- Hamandawana, H. (2008). Monitoring transitions in vegetation cover to detect long-term trends in ground water resources. 2008 UNESCO UCI Groundwater Conference Proceedings.
- Huffman, G.J., R.F. Adler, M. Morrissey, D.T. Bolvin, S. Curtis, R. Joyce, B McGavock, J. Susskind, 2001: Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations. *Journal of hydrometeorology*, 2(1), 36-50.
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., ... & Stocker, E. F. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of hydrometeorology*, 8(1), 38-55.
- Huete, A., Justice, C., & Liu, H. (1994). Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of Environment*, 49(3), 224-234.
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote sensing of environment*, 83(1-2), 195-213.
- Ingram, H. A. P. (1987). Ecohydrology of Scottish peatlands. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, 78(4), 287-296.
- Jin, C., Xiao, X., Merbold, L., Arneeth, A., Veenendaal, E., & Kutsch, W. L. (2013). Phenology and gross primary production of two dominant savanna woodland ecosystems in Southern Africa. *Remote Sensing of Environment*, 135, 189-201.
- Kling, H., Stanzel, P., & Preishuber, M. (2014). Impact modelling of water resources development and climate scenarios on Zambezi River discharge. *Journal of Hydrology: Regional Studies*, 1, 17-43.

- Kurugundla, C. N., Parida, B. P., & Buru, J. C. (2018). Revisiting Hydrology of Lake Ngami in Botswana. *Hydrol. Current Res.*, 9 (301), 2.
- Lillesand, T., Kiefer, R. W., & Chipman, J. (2014). Remote sensing and image interpretation. John Wiley & Sons.
- López E., Bocco G., Mendoza M., Duhau E.; 2001. Predicting land-cover and land-use change in the urban fringe: A case in Morelia city, Mexico. *Landscape and Urban Planning*, 55, 271 – 285.
- Ma, X., Huete, A., Yu, Q., Coupe, N. R., Davies, K., Broich, M., ... & Boulain, N. (2013). Spatial patterns and temporal dynamics in savanna vegetation phenology across the North Australian Tropical Transect. *Remote sensing of Environment*, 139, 97-115.
- Mbaiwa, J. E. (2005). Enclave tourism and its socio-economic impacts in the Okavango Delta, Botswana. *Tourism Management*, 26(2), 157-172.
- Meier, S. D., Atekwana, E. A., Molwalefhe, L., & Atekwana, E. A. (2015). Processes that control water chemistry and stable isotopic composition during the refilling of Lake Ngami in semiarid northwest Botswana. *Journal of Hydrology*, 527, 420-432.
- Milzow C., Kgotlhang L., Bauer-Gottwein P., Meier P., Kinzelbach W.; 2009. Regional review: the hydrology of the Okavango Delta, Botswana – processes, data and modelling. *Hydrogeology Journal*, 17, 1297 – 1328.
- Mishra, N. B., Crews, K. A., & Neuenschwander, A. L. (2012). Sensitivity of EVI-based harmonic regression to temporal resolution in the lower Okavango Delta. *International journal of remote sensing*, 33(24), 7703-7726.
- Mosimanyana, E., Kurugundla, C. N., & Murray-Hudson, M. (2016). Analysis of the hydrological processes of a semi-arid region lake, the case of Lake Ngami, Botswana. In Botswana Symposium on Wetlands and Wildlife.

- Munyati, C., & Mboweni, G. (2013). Variation in NDVI values with change in spatial resolution for semi-arid savanna vegetation: a case study in northwestern South Africa. *International journal of remote sensing*, 34(7), 2253-2267.
- Murray-Hudson, M., Wolski, P., Cassidy, L., Brown, M. T., Thito, K., Kashe, K., & Mosimanyana, E. (2015). Remote Sensing-derived hydroperiod as a predictor of floodplain vegetation composition. *Wetlands ecology and management*, 23(4), 603-616.
- Okavango Research Institute. 2018. Okavango Data: Monitoring and forecasting. Maun, University of Botswana. Retrieved from: <http://okavangodata.ub.bw/ori/>
- Pedzisai, E., & Gondo, R. (2016). Spatial policy in natural resources governance. *International Journal of Environmental Protection and Policy*, 4(5), 141-154.
- 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.
- QGIS Development Team. (2016). QGIS geographic information system. Open Source Geospatial Foundation Project.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Rodriguez-Iturbe I., 2000. Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamics. *Water Resources Research*, 36, 3 – 9.
- Schwinning, S., Sala, O. E., Loik, M. E., & Ehleringer, J. R. (2004). Thresholds, memory, and seasonality: understanding pulse dynamics in arid/semi-arid ecosystems. *Oecologia*, 14, 191 – 193.
- Shaw P (1985) Late quaternary landforms and environmental change in Northwest Botswana: the evidence of Lake Ngami and the Mababe depression. *Transactions of the Institute of British Geographers*, 333–346



- UNDP. 2014a. Mainstreaming SLM in Rangeland Areas of Ngamiland District Landscapes for Improved Livelihoods. UNDP. UN.
- UNESCO (1971) Convention on Wetlands of International Importance Especially as Waterfowl Habitat. Ramsar, Iran, 2 February 1971, UN Treaty Series No. 14583. As amended by the Paris Protocol, 3 December 1982, and Regina Amendments, 28 May 1987.
- von Hardenberg, J., Meron, E., Shachak, M., & Zarmi, Y. (2001). Diversity of vegetation patterns and desertification. *Physical Review Letters*, 87(19), 198101.
- Walker, J. J., De Beurs, K. M., & Wynne, R. H. (2014). Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and Landsat data. *Remote Sensing of Environment*, 144, 85-97.
- White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global biogeochemical cycles*, 11(2), 217-234.
- Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.
- Wickham, H., & Henry, L. (2017). *tidyr: Easily Tidy Data with spread () and gather () Functions*. R package version 0.6, 3.
- Wolski, P., & Murray-Hudson, M. (2006). Recent changes in flooding in the Xudum distributary of the Okavango Delta and Lake Ngami, Botswana. *South African Journal of Science*, 102(3-4), 173-176.
- Yamazaki, D., Trigg, M. A., & Ikeshima, D. (2015). Development of a global~ 90 m water body map using multi-temporal Landsat images. *Remote Sensing of Environment*, 171, 337-351.
- Zalewski, M. (2014). Ecohydrology and hydrologic engineering: regulation of hydrology-biota interactions for sustainability. *Journal of Hydrologic Engineering*, 20(1), A4014012.
- Zhang, S., Ye, Z., Chen, Y., & Xu, Y. (2017). Vegetation responses to an ecological water conveyance project in the lower reaches of the Heihe River basin. *Ecohydrology*, 10(6), e1866.

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*, 9(4), 256.

## Appendix I: Google Earth Engine Scripts

### Appendix I.I: Global Surface Water (GSW)

Script to process and extract monthly extent of the Lake Ngami extent from the Global Surface Water data set. The script is also available on:

[https://code.earthengine.google.com/gsw\\_processing](https://code.earthengine.google.com/gsw_processing)

```

////////////////////////////////////
//          SCRIPT FOR SURFACE WATER ANALYSIS IN THE OKAVANGO          //
//          Eduardo Garcia Bendito (10/2018)                          //
////////////////////////////////////

// Create start and end date of the analysis
var startDate = '2000-01-01';
var endDate = '2013-12-31';

////////////////////////////////////
////////////////////////////////////

// Create a polygon of the maximum extent of water of Lake Ngami
// Load Global Surface Water Dataset
var gsw = ee.Image('JRC/GSW1_0/GlobalSurfaceWater');
// select the 'Maximum Water Extent' layer
var max = gsw.select('max_extent');
// create a vector from the 'Maximum Water Extent'
var maxExtent = max.reduceToVectors({
  geometry: lake,
  crs: gsw.projection(),
  scale: 10,
  geometryType: 'polygon',
  eightConnected: true,
  labelProperty: 'max'
});

// export the FeatureCollection to a SHP file.
Export.table.toDrive({
  collection: maxExtent,
  description: 'Lake_Ngami',
  fileFormat: 'SHP'
});

// The buffer feature is ingested into GEE after exporting the FeatureCollection
// 'maxExtent' as a .shp and processing in QGIS. Initially the vector file (.shp) has
// many features. The largest feature representing the maximum extent of the lake is
// selected and a "buffer" is applied on this element using a distance of 5 km (in the
// QGIS dialog of the "Buffer" tool, 'Distance' = 0.05).

////////////////////////////////////
////////////////////////////////////

// Load the monthly Global Surface Water dataset and

```

```
// add bands of monthly water detected, no water, and no data
var monthlyHistory = ee.ImageCollection("JRC/GSW1_0/MonthlyHistory")
  .filterDate(startDate, endDate)
  .map(function(image){
    image = image.addBands(image.select(["water"],["yes"]).eq(2));
    image = image.addBands(image.select(["water"],["no"]).eq(1));
    image = image.addBands(image.select(["water"],["noData"]).eq(0));
    return image;
  });

////////////////////////////////////
////////////////////////////////////

//MAPPING SECTION
Map.centerObject(middle, 9);
//print(monthlyHistory, 'months');
var vizParams1 = {
  bands: ['yes'],
  min: 0,
  max: 1,
  palette: ['white', 'blue']
};
var vizParams2 = {
  bands: ['no'],
  min: 0,
  max: 1,
  palette: ['white', 'red']
};
var vizParams3 = {
  bands: ['noData'],
  min: 0,
  max: 1,
  palette: ['white', 'black']
};
Map.addLayer(ee.Image(monthlyHistory.toList(1,11).get(0)),vizParams1, "Surface
Water");
Map.addLayer(ee.Image(monthlyHistory.toList(1,11).get(0)),vizParams2, "No Water",
false);
Map.addLayer(ee.Image(monthlyHistory.toList(1,11).get(0)),vizParams3, "No Data",
false);
Map.addLayer(lakeNgami,{}, "Maximum Water Extent Lake Ngami");

////////////////////////////////////
////////////////////////////////////

// SURFACE WATER ANALYSIS
// Calculate the area of water
var water = monthlyHistory.map(function(image){
  var waterAreas = image.mask(image.select(["yes"]).eq(1));
  image = image.addBands(ee.ImageCollection([
    waterAreas.select(["yes"]).toInt()).reduce(ee.Reducer.max()));
  //get the water presence as n° of pixels
  image = image.addBands(ee.ImageCollection([
    waterAreas.select(["yes"]).toInt(),
  ]).reduce(ee.Reducer.max()).rename(["water"]));
});
```

```

    //get the areas of water in km-2
    return
image.select(["yes"]).multiply(ee.Image.pixelArea().divide(1000000)).copyProperties(i
mage);
});

//Calculate the area of No Data
var noData = monthlyHistory.map(function(image){
  var noData = image.mask(image.select(["noData"]).eq(1));
  image = image.addBands(ee.ImageCollection([
    noData.select(["noData"]).toInt()]).reduce(ee.Reducer.max()),
    null, true);
  //get No Data as n° of pixels
  image = image.addBands(ee.ImageCollection([
    noData.select(["noData"]).toInt(),
  ]).reduce(ee.Reducer.max()).rename(["NoData"]), null, true);
  //get the areas of water in km-2
  return
image.select(["noData"]).multiply(ee.Image.pixelArea().divide(1000000)).copyPropertie
s(image);
});

//Calculate the area of both water and no-data
var both = monthlyHistory.map(function(image){
  var water = image.mask(image.select(["yes"]).eq(1));
  var noWater = image.mask(image.select(["no"]).eq(1));
  var noData = image.mask(image.select(["noData"]).eq(1));
  image = image.addBands(ee.ImageCollection([
    water.select(["yes"]).toInt(),
    noWater.select(["no"]).toInt(),
    noData.select(["noData"]).toInt()]).reduce(ee.Reducer.max()));
  //get the water presence as n° of pixels
  image = image.addBands(ee.ImageCollection([
    water.select(["yes"]).toInt(),
  ]).reduce(ee.Reducer.max()).rename(["water"]))
  .addBands(ee.ImageCollection([
    noWater.select(["no"]).toInt()
  ]).reduce(ee.Reducer.max()).rename(["noWater"]))
  .addBands(ee.ImageCollection([
    noData.select(["noData"]).toInt()
  ]).reduce(ee.Reducer.max()).rename(["noData"]));
  //get the areas of water in km-2
  return image.select(["yes", "no",
"noData"]).multiply(ee.Image.pixelArea().divide(1000000)).copyProperties(image);
});

//get the total areas for the user-defined polygon
var trendWater = ee.FeatureCollection(water.map(function(image){
  var water = ee.FeatureCollection(image.reduceRegions(lakeNgami, ee.Reducer.sum(),
30).first());
  water = water.set("year", image.get("year"));
  water = water.set("month", image.get("month"));
  return water;
}));

```

```
var trendNoData = ee.FeatureCollection(noData.map(function(image){
  var feature = ee.FeatureCollection(image.reduceRegions(lakeNgami, ee.Reducer.sum(),
30).first());
  feature = feature.set("year", image.get("year"));
  feature = feature.set("month", image.get("month"));
  return feature;
}));

var trendboth = ee.FeatureCollection(both.map(function(image){
  var feature = ee.FeatureCollection(image.reduceRegions(lakeNgami, ee.Reducer.sum(),
30).first());
  feature = feature.set("year", image.get("year"));
  feature = feature.set("month", image.get("month"));
  return feature;
}));

////////////////////////////////////
////////////////////////////////////

//CHART SECTION
var options1 = {
  legend: { position: 'top', maxLines: 3 },
  vAxis: {title: 'Area Km-2'},
  hAxis: {title: 'Month', format: '####'},
  bar: { groupWidth: '75%' },
  isStacked: true,
  title: 'Lake Ngami area 1984-2015'
};
var options2 = {
  legend: { position: 'top', maxLines: 3 },
  vAxis: {title: 'Area Km-2'},
  hAxis: {title: 'Month', format: '####'},
  bar: { groupWidth: '75%' },
  isStacked: true,
  title: 'No Data area 1984-2015'
};
var options3 = {
  legend: { position: 'top', maxLines: 3 },
  vAxis: {title: 'Area Km-2', viewWindow: { max: 250 }},
  hAxis: {title: 'Month', format: '####'},
  bar: { groupWidth: '50%' },
  isStacked: true,
  title: 'Lake Ngami water presence 2000-2013'
};

print(ui.Chart.feature.byFeature(trendboth, "system:index",
["yes","no","noData"]).setOptions(options3).setChartType("ColumnChart"));
```

## Appendix I.II: Climate Hazards InfraRed Precipitation with Stations (CHIRPS)

Script to process and extract monthly precipitation around Lake Ngami from the Climate Hazards InfraRed Precipitation with Stations (CHIRPS) data set. The script is also available on: [https://code.earthengine.google.com/chirps\\_processing](https://code.earthengine.google.com/chirps_processing)

```

////////////////////////////////////
//          SCRIPT FOR CHIRPS ANALYSIS IN LAKE NGAMI          //
//          Eduardo Garcia Bendito (10/2018)                  //
////////////////////////////////////

// Create the time window from 1st Jan 2000 until 31 Dec 2012
var month_mean = ee.List.sequence(0, 13*12).map(function(n) { // .sequence: number of
months from starting year to present
  var start = ee.Date('2000-01-01').advance(n, 'month'); // Starting date
  var end = start.advance(1, 'month'); // Step by each iteration
// Filter CHIRPS for the time-window selected
  return ee.ImageCollection("UCSB-CHG/CHIRPS/DAILY")
    .filterDate(start, end)
    .sum()
    //.mean()
    .set('system:time_start', start.millis());
});

var collection = ee.ImageCollection(month_mean);

// Create an array of the time series in the geometry of Lake Ngami Study Area
var area_name = fullSA.aggregate_array('name').getInfo();
var title = 'CHIRPS [mm/month] for Buffered area Lake Ngami';

// Plot the time series
var TimeSeries = ui.Chart.image.seriesByRegion({
  imageCollection: collection,
  //regions: fullSA,
  //regions: multi_bu,
  regions: multi_bu,
  reducer: ee.Reducer.mean(),
  scale: 1000,
  xProperty: 'system:time_start',
  seriesProperty: 'label'
}).setChartType('ColumnChart')
.setOptions({
  title: title,
  vAxis: {title: '[mm/year]',
  //viewWindow: {max: 200}
  },
  lineWidth: 1,
  pointSize: 1,
});

print('Precipitation time-series of study area:', TimeSeries);

```

```
//Print the maps  
Map.addLayer(ee.Image(collection.toList(1,1).get(0)),{bands: ['precipitation'],min:  
0, max: 200, palette: ['white', 'blue']}, "Jan 2000");  
Map.addLayer(multi_bu, {palette: ['red']}, 'Study Area');
```



## Appendix I.III: MODIS 16-Day composite EVI

Script to process and extract EVI around Lake Ngami from MODIS 16-Day composite data set.

The script is also available on: [https://code.earthengine.google.com/modis\\_processing](https://code.earthengine.google.com/modis_processing)

```

////////////////////////////////////
//          SCRIPT FOR MODIS EVI AROUND LAKE NGAMI          //
//          Eduardo Garcia Bendito (10/2018)                //
////////////////////////////////////

// Create start and end date of the analysis
var startDate = '2000-02-20';
var endDate = '2012-12-31';

// Collection of buffers
var classes = new ee.FeatureCollection([b_1km,b_2km,b_3km,b_4km,b_5km,
                                       b_6km,b_7km,b_8km,b_9km,b_10km]);
var classes_flatten = classes.flatten();
var regions = classes_flatten;

////////////////////////////////////
//LOAD THE DATASET
// Load the MODIS Combined 16-Day EVI Collection
var modis16Day = ee.ImageCollection('MODIS/MCD43A4_006_EVI')
// Filter through the dates
  .filterDate(startDate, endDate);

////////////////////////////////////

//get the mean EVI for buffer 1
var eviMOD1 = ee.FeatureCollection(modis16Day.select("EVI").map(function(image){
  var mean = ee.Feature(image.reduceRegions({
    collection: b_1km,
    reducer: ee.Reducer.mean(),
    scale:250
  })
  .first());
  mean = mean.set("system:index", image.get("system:index"));
  return mean;
}));
var eviMOD1 = eviMOD1.select(['mean']).select(['.*'],null,false);

// ... repeat for all buffers

//get the mean EVI for buffer 10
var eviMOD10 = ee.FeatureCollection(modis16Day.select("EVI").map(function(image){
  var mean = ee.Feature(image.reduceRegions({
    collection: b_10km,
    reducer: ee.Reducer.mean(),
    scale:250
  })
  .first());
  mean = mean.set("system:index", image.get("system:index"));
  return mean;
}));

```

```
    }));
    var eviMOD10 = eviMOD10.select(['mean']).select(['.*'],null,false);
    print(eviMOD10);

    // Export time series of each buffer

    Export.table.toDrive({
      collection: eviMOD11,
      description: 'EVI_1km_2000-2012',
      fileFormat: 'CSV'
    });

    Export.table.toDrive({
      collection: eviMOD10,
      description: 'EVI_10km_2000-2012',
      fileFormat: 'CSV'
    });

    var optionsEVI = {
      legend: { position: 'top'},
      vAxis: {title: 'Mean EVI'},
      hAxis: {title: 'Month', format: '####'},
      title: 'Mean EVI 2000-2015'
    };

    print(ui.Chart.feature.byFeature(eviMOD10, "system:index",
    ["mean"]).setOptions(optionsEVI).setChartType("LineChart"));

    var chart = ui.Chart.image.seriesByRegion(
      modis16Day, regions, ee.Reducer.mean(), 'EVI', 250, 'system:index', 'distance')
      .setChartType('LineChart')
      .setOptions({
        vAxis: {title: 'EVI'},
      });

    // Display.
    print(chart);
```

## Appendix II: Global Surface Water uncertainty

The Global Surface Water data set includes three classes for each pixel: Water, No Water or No Data. For each month, the amount of uncertainty was assessed for the mapping of open water in Lake Ngami. The results indicate that on average, there is a 27 % uncertainty in this data set throughout the entire time-series in Lake Ngami.

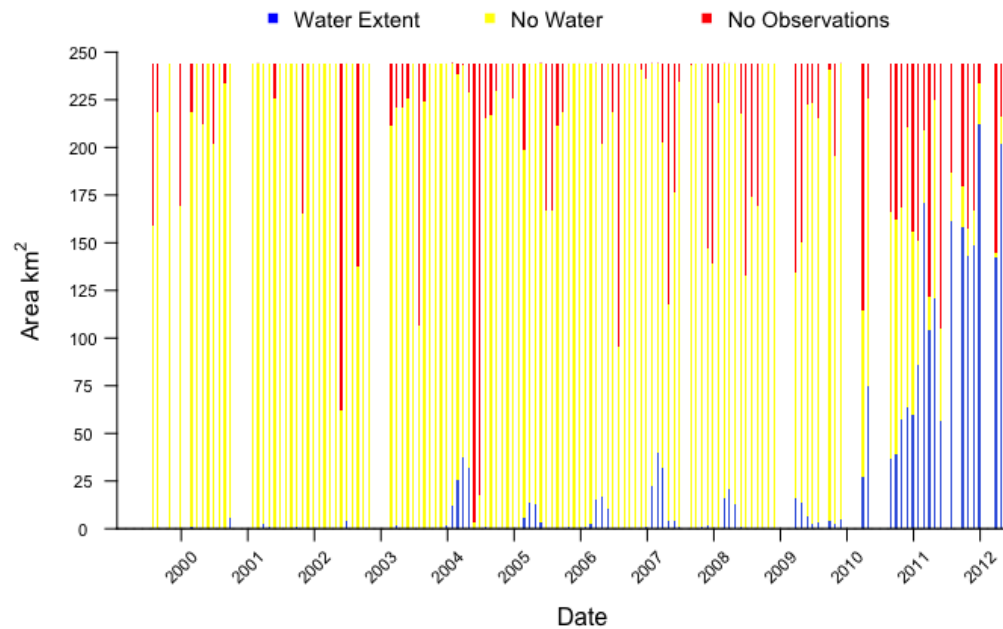


Figure II: Monthly uncertainty analysis of the measurements in the maximum extent of Lake Ngami area coverage (km<sup>2</sup>) between 2000 and 2012 derived from the Global Surface Water (GSW) data set. The red columns indicate 'No Data' for that month. That area could be Water or No Water class.



## Appendix III: Annual phenology and precipitation (2000 - 2012)

Annual dynamics of vegetation phenology and precipitation starting in July for each year of the study period (2000 - 2012). The phenological phases including the minimum and maximum photosynthetic activity are also included (Min and Max EVI), indicating the length of the greening season. These data are the mean composite aggregation of the each of the 1 km buffer stripes EVI time series, thus representing the average vegetation dynamics around Lake Ngami.

