Assessment of nature visitation in Lake Tana Biosphere Reserve through crowdsourced data using social media

MARTHA WELDEMICHAEL DEREJE
June 2021

SUPERVISORS:
Prof. Dr. Louise (Wieteke) Willemen
Dr. Frank O. Ostermann

ADVISOR
Dr. Evangelia (Valia) Drakou

Assessment of nature visitation in Lake Tana Biosphere Reserve through crowdsourced data using social media

MARTHA WELDEMICHAEL DEREJE
Enschede, The Netherlands, June 2021

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

SUPERVISORS:
Prof. Dr. Louise (Wieteke) Willemen
Dr. Frank O. Ostermann

ADVISOR
Dr. Evangelia (Valia) Drakou

THESIS ASSESSMENT BOARD:
Dr.ir. C.A.J.M. de Bie (Chair)
Dr. Katja Egorova (External examiner)

DISCLAIMER

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

ABSTRACT

Cultural Ecosystem Services (CES) contribute to cultural identity, livelihoods, and even survival, but are hard to measure and quantify due to their intangible and non-consumptive attributes. However, rapid growth in mobile network connections and the usage of social media has resulted in huge amounts of crowdsourced data. For measuring and mapping CES, geolocated and timestamped social media content are becoming more popular. In this study, social media platforms that are used for nature visitation assessment have been identified using a systematic literature review. The review of the studies reveals that social media platforms contain geolocation and timestamps, allowing for analysis of the spatial and temporal pattern of nature visitation. In this study, Flickr and Twitter are evaluated for their ability to describe nature visitation in space and time as they are open to access data. The locations and the time of photographs and tweets were used to assess the spatial and temporal distribution of nature visitation in the Lake Tana Biosphere Reserve. The spatial visitation pattern of the Lake Tana Biosphere Reserve was identified by quantifying the proxy Photo User Day (PUD) and Tweet User Day (TUD). This research shows that most PUDs and TUDs are clustered in the south (around Bahir Dar) and northeast part of the area and with some PUDs and TUDs in Lake Tana. The research also shows that most PUDs are close to the road and on the way to Blue Nile Falls (Tis-Isat Falls). The January month, that has two colourful festivals, has a higher number in both PUD and TUD. Assessing the spatial and temporal pattern of nature visitation offers to understand which elements of nature attract people to locations around the Lake Tana Biosphere Reserve and when people came.

ACKNOWLEDGEMENTS

Praises and thanks to the God, the Almighty, for his unlimited blessings throughout my life.

I would like to acknowledge the Orange Knowledge Program (OKP) for offering me the fellowship that covers overall costs related to study and my stay in the Netherlands.

I would like to express my deep and sincere gratitude to my first supervisor Prof. Dr. Louise (Wieteke) Willemen for giving me invaluable guidance throughout this research. Her guidance in reviewing my work, providing feedback and ideas, and her consistency in enquiring about my progress encourage me to keep working hard. I thank her for being patient with me in assisting me. My gratitude also goes to my second supervisors Dr. Frank O. Ostermann and Dr. Evangelia (Valia) Drakou for their encouragement and constructive comments. I want to thank Tegegne Molla Sitotaw for providing me data and information about the study area.

My warm and heartfelt thanks go to my family for their love, prayers, advice and hope they had given to me.This thesis would not have been possible without that hope. Finally, Devaye. You were always there, whenever, and wherever that was. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Humans rely on the environment they live in and that encompasses them. They are dependent on the earth's ecosystems and the services that these provide. So far, several definitions of Ecosystem Services (ES) exist. ES are defined as "the benefits that people get from ecosystems" or can be described as "direct and indirect contributions of ecosystems to human well-being" (Crossman et al., 2013). There are different types of ES these include provisioning services (products obtained from ecosystems), regulating services (such as climate, water quality, clean air processes), and cultural services. Cultural ecosystem services (CES) are defined as "the non-material benefits people obtain from ecosystems through spiritual enrichment, cognitive development, reflection, recreation, and aesthetic experiences" (Müller et al., 2019).

CES are potential motivators and incentives for people to protect their environment (Müller et al., 2019). They contribute to human well-being, public health, and psychological experiences at individual and collective levels (Hirons et al., 2016). When destroyed, these services are significantly harder to supplant than other ecosystem services (Helka, 2016). CES are often essential for cultural identity, livelihoods, and even survival, but it is hard to measure and quantify them because of their elusive and non-consumptive properties (Hirons et al., 2016). However, data that are collected traditionally like household surveys, interviews, national statistics data, and the huge amount of crowdsourced data that are generated from social media offers an opportunity to measure and quantify CES.

Several previous studies have been conducted to assess CES based on survey data and interviews (Codato et al., 2017; Plieninger et al., 2013). Also, CES can be assessed and mapped using participatory mapping, and public participation geographic information system (PPGIS) can be used for collecting and using non-expert spatial information (Canedoli et al., 2017). Individual surveys and interviews are important in getting ideas and people's feelings about the services and they support the participation of stakeholders in CES valuation (Delgado-Aguilar et al., 2017). Spatial and temporal information is needed to measure and quantify cultural ecosystem services because supply and demand for ecosystem services vary over space and time (Troy & Wilson, 2006). In the case of surveys and interviews, they have a limited scope on time and space, but there is an alternative method like crowdsourced data to widen the scope.

Currently, crowdsourced data has gained significant recognition relative to field data for the rapid growth of its usage and data generation, such as geotagged images and tweets is estimated to be 2.5 exabytes per day (Kim et al., 2019). Vast amounts of photographs which are tagged with exact information about the time and area where they were taken have been produced with the integration of smartphones and Global Navigation Satellite System (GNSS) technology like GPS, Galileo, GLONASS and other regional systems in different portable devices (Havinga et al., 2020). Geotagged photographs from social media (Instagram, Flickr, Twitter, Panoramio, or Facebook) can be used to map CES, such as nature-based recreation patterns, value recreational ecosystem services, investigate how recreational benefits are affected by changes in ecosystem quality (Havinga et al., 2020).

Natural ecosystems, islands, and wetlands provide habitat for a diverse range of animal and plant species. Lake Tana Biosphere Reserve is rich in biological diversity. In Ethiopia, 50 % of highland areas are covered by Afromontane forests, and Lake Tana Biosphere Reserve is in the part of Eastern Afromontane biodiversity hotspot (Worku, 2017). Lake Tana Biosphere Reserve was established to safeguard the

enormous biological and cultural richness of the area around Lake Tana, while also promoting sustainable economic growth and land use(Getnet Fetene, 2017). The biosphere reserve comprises mainly of agricultural land, cropland, water body, church forests, and a city. The majority of the Lake Tana cultural attractions are churches and monasteries on the islands and peninsulas on the Lake. Forests other than church areas have been destroyed over centuries, and this is because of grazing activities, tree harvesting, farming, and other activities. However, small patches of forests that are around the churches contribute to the restoration, biodiversity conservation and provide many other societal, scientific, and economic benefits (Alemayu Wassie Eshete, 2007). Forests surrounding the churches provide crucial economic, ecological, and recreational opportunities also are used for social gatherings and religious ceremonies (high visitation rate during holiday celebrations) (Kassahun & Bender, 2019).

1.1. Problem statement

Afromontane forest fragments in Ethiopia and wetlands are known for their ecological relevance; however, the benefits they supply from non-material services are not well-understood and recognized. Furthermore, natural areas are not equally important to people over space and time. Some areas are preferred over other areas because they are close to the road (accessibility) or are more important because of their land cover type. Also, natural areas can be more attractive at a specific time because of seasonal changes in nature and people's travel preferences. This research will focus on the use of crowdsourced information on nature visitation through geotagged and timestamped posts, as a proxy for a human appreciation of nature in the Lake Tana Biosphere Reserve. Because of limited and inconsistent data, measuring and quantifying this non-material service in the area is not well investigated. Earlier studies on nature visitation used qualitative methods like surveys and interviews within small study areas (Alemayu Wassie Eshete, 2007; Wondie, 2018). However, these methods are time-consuming, laborious, and costly. Social media data can be an alternative because they are often openly available, have good spatial coverage as well as we can use them to overcome data limitations. There are differences between platforms, and social media data appears to work differently, so we should select the platforms with caution. The most commonly used social media platforms on which people post messages of places visited are Facebook, Instagram, Flickr and Twitter. Most research that used social media data to analyze nature visitation focused on a single social media platform so far and as a result, there is a lack of formal evaluation of various social media platforms to assess nature visitation (Tenkanen et al., 2017). Therefore, it is important to compare and contrast content structure and attributes of commonly used social media platforms and the implications for assessing nature visitation. Working with the social media platforms is important for tourism mangers and planners to combine the data with the traditional survey data so that to inform sustainable tourism development plans and policy decisions.

1.2. Research Objectives and questions

The research's overall objective is to use crowdsourced data to assess nature visitation on Lake Tana Biosphere Reserve in Ethiopia. To address the overall objective the following specific objectives and research questions are formulated.

1. Compare and contrast content structure and attributes of Flickr and Twitter platforms to assess nature visitation.
    1.1. What are the key characteristics of structure, attributes, and content of these platforms for assessing visitation?
    1.2. What are the differences and similarities in structure, content, and attributes between the two platforms?
2. To assess the spatial distribution of nature-visitation in Lake Tana Biosphere Reserve based on the two social media platforms.
    2.1. What is the spatial pattern of locations being visited by people?

2.2. What social and biophysical characteristics of the visited locations can explain the pattern?
2.3. What is the difference and similarity in spatial pattern between the two platforms?

3. To assess the temporal distribution of nature-visitation in Lake Tana Biosphere Reserve.
3.1. What is the temporal pattern of the visited locations within a month?
3.2. What is the difference and similarity in temporal pattern between the platforms?

The results of the first objective are given in Chapter 3, while 2 and 3 are answered in Chapter 4.

2. STUDY AREA

2.1. Study Area

The study is conducted in Lake Tana Biosphere Reserve (Figure 1), which is a protected area located in the north-western part of Ethiopia. In 2011 a feasibility study was conducted by the Michael Soccow Foundation in collaboration with Nature and Biodiversity Conservation Union (NABU) to assess the Lake Tana region's potential as a biosphere reserve (Zur Heide, 2012). The feasibility study shows that the Lake Tana Region fulfils the prerequisites to comply with the Biosphere Reserve designation criteria of United Nations Educational, Scientific and Cultural Organization (UNESCO). The biosphere reserve comprises Lake Tana, one of the foundations of the Blue Nile River, which provides important ecosystem services, and the church forest ecosystem provides many valuable cultural services to human beings (Kassahun & Bender, 2019). The total area covers 6972 square km, of which 3042 square km represents the aquatic surface and 3930 square km terrestrial (Kalmbach, 2017). The Biosphere Reserve comprises more than 300 bird species, 21 endemic fishes, and >180 woody plants (Kalmbach, 2017). More than 2 million people live in the Biosphere Reserve, with the most populated people in the region being the Amhara people (Bires & Raj, 2019). An enormous heterogeneity of land uses, and natural ecosystems characterize the region. Agriculture, fishing, national and international tourism (religious and recreational), and sand mining are the major economic activities in the Biosphere Reserve (Kalmbach, 2017).

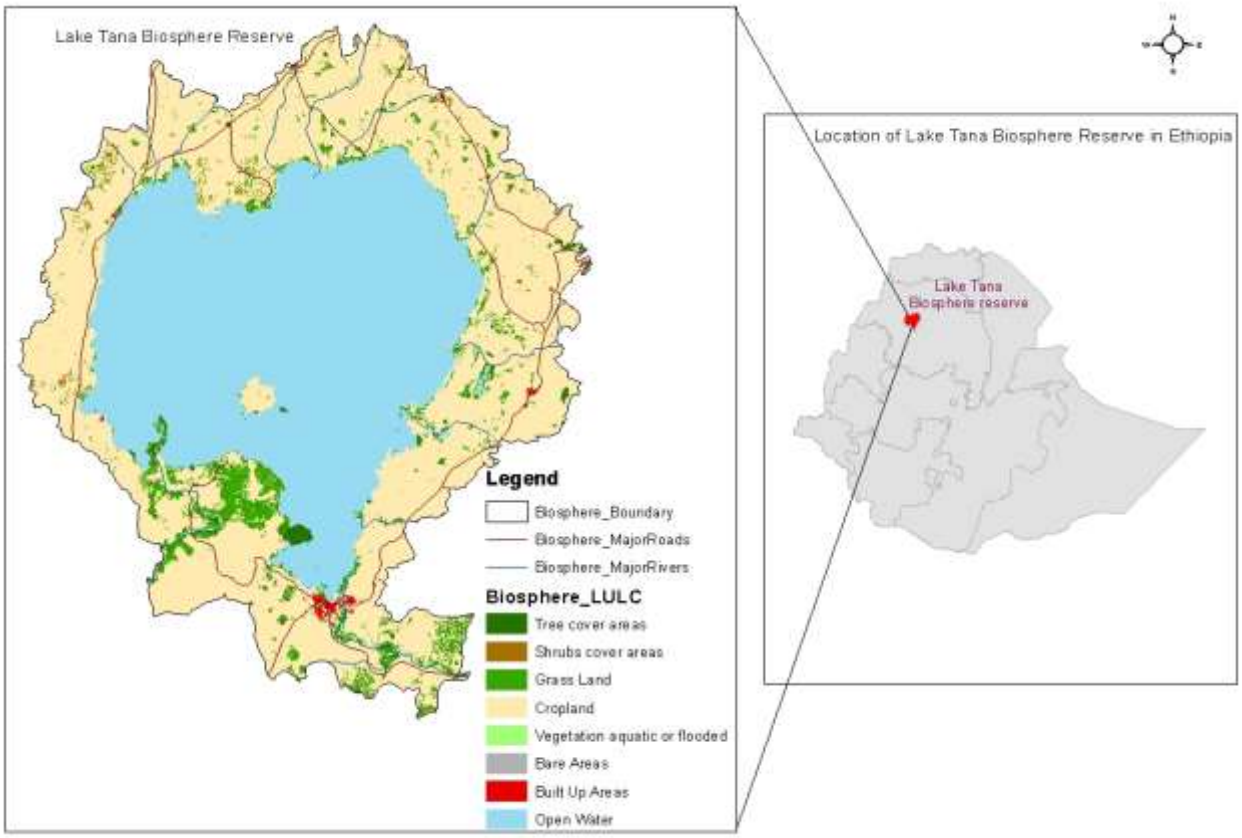

Figure 1. Study Area (Land cover: ESA Climate Change Initiative - Land Cover project 2017)

3. MATERIALS AND METHODS

3.1. Method Workflow

Several activities were performed to address the objective of this study: social media platform selections, data acquisition from selected social media platforms, cleaning and pre-processing the data, data analysis, mapping of geotagged photos, analysing the data based on land cover type and accessibility. The overall workflow of the study is indicated in figure 2 below.

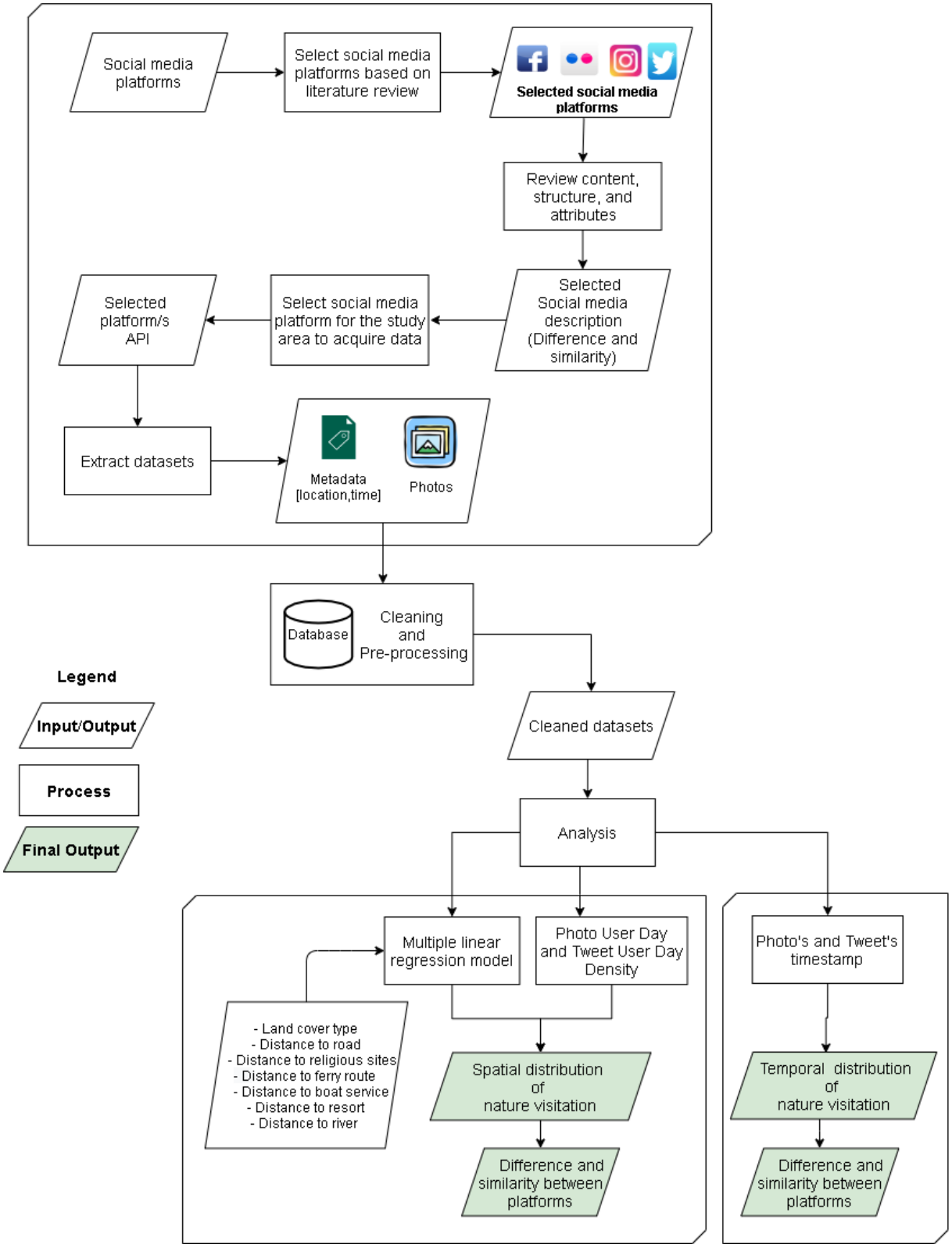

Figure 2. Method Workflow

3.2. Literature Review

3.2.1. Review approach

To analyse research that uses social media platforms for assessment of nature visitation a review was conducted using a systematic quantitative literature review. To clarify the process of selecting and rejecting articles for the literature review, the PRISMA statement idea is used (Moher et al., 2009). Initial paper selection was achieved through searches in two databases (Scopus and Web of Science) using search terms (("social media") AND (tourism OR visitation) AND (nature OR "natural areas")) from the title, abstract and keywords. Additional papers are obtained from supervisor's suggestions. The method that was used to filter the results in each step and final results is depicted in figure 3. All results that are published in English and within the most recent 10 years (2012-2021) were exported and processed further. After removing the duplicates, the titles and abstract of the remaining publication were assessed and publications related to chemistry, medicine, nursing, and others that did not use social media platforms were removed. From the remaining publications, papers that did not focus on nature visitation and did not use social media platforms to assess their work were removed based on their abstract. Finally, the remaining paper were assessed quantitatively and qualitatively based on 1) How many platforms are used? 2) What kind of platforms they used for their study? 3) What kind of data attributes are used in their study? 4) What are the characteristics of the platforms they used in their study? The final 12 results do not combine other data sources like interview, survey data etc., on their study, they only use social media platforms. The results of the papers help to learn what attributes and platforms is needed to assess nature visitation and help to be able to select a platform and use the right attributes for the study.

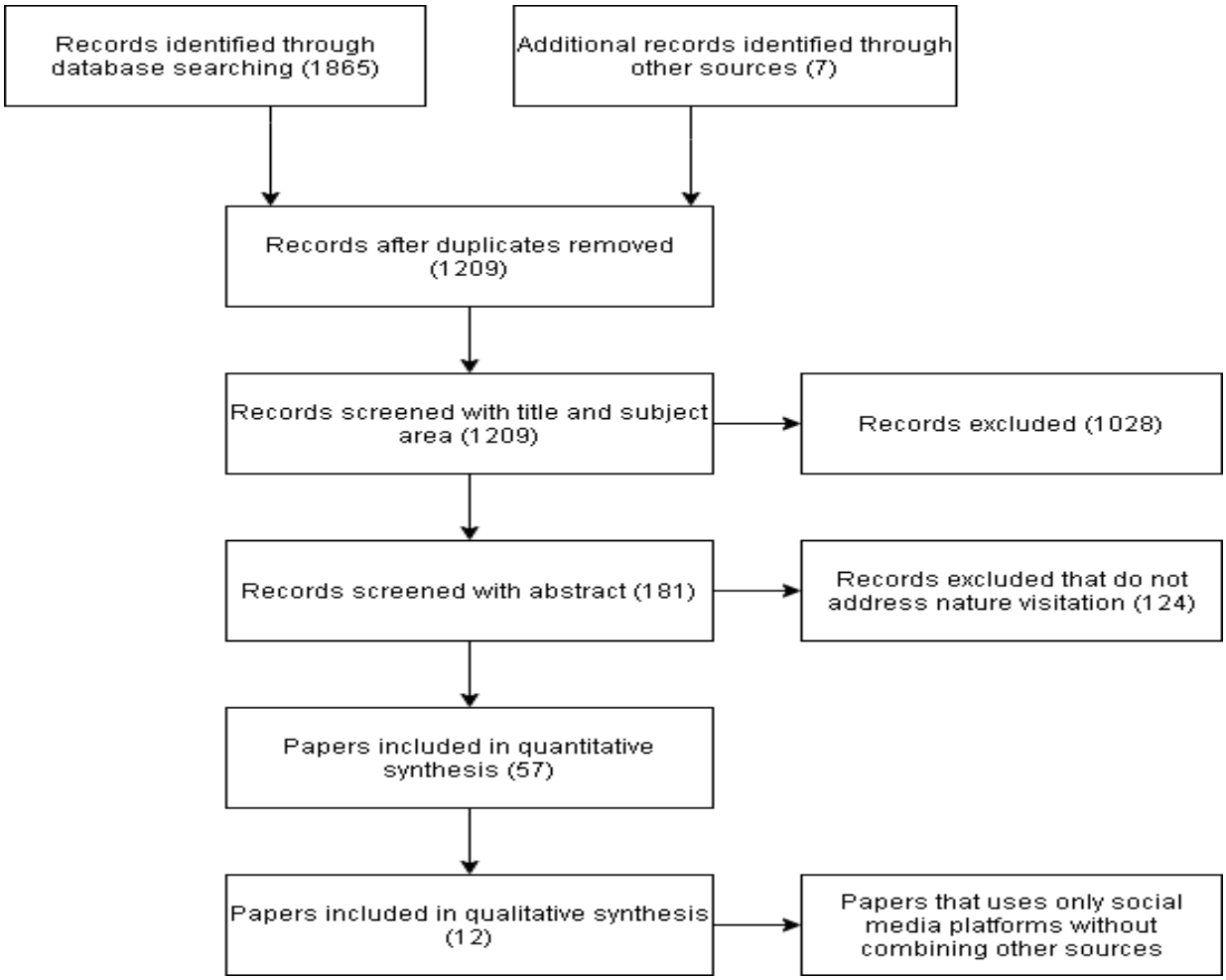

Figure 3. Literature review process

From the literature review it was found that most papers use a single social media platform for their study, and some use two or three platforms and do a comparison. Almost every publication included some form of spatial analysis and there are articles that look at image content, temporal, and text analysis. In data analysis, spatial and temporal analysis using social media platforms were frequently used method. Some of the publications that used spatial analysis looked at aspects that could influence visitation, such as the environment, destination infrastructure, or visitor social-demographic characteristics. Other studies used picture geolocation and park attractiveness to estimate visitor flows in natural areas. The origin of visitors (for example, where they came from) was an important topic in several articles. In some articles, the distance to the visitor's home to the sites were assessed. Studies that looked at temporal patterns of visitors looked at yearly variation. Several studies looked at seasonal and/or monthly patterns. Another study looked at the temporal variance in leisure activities in a well-known national park. Some publications looked at predicting visitation, with some of them focusing on the impact of changing environmental conditions on visitor numbers. So, the review of the studies reveals that social media platforms contain geolocation and timestamps allowing for analysis of spatial and temporal pattern of nature visitation.

3.2.2. Assessment of social media platforms

As this research is to learn about visitation meaning the following aspect, the where? and when? (Figure 4) are important so looking at the data these are the most important things. Before collecting the data, the platforms need to be assessed and to be able to assess the suitability of social media platforms it is evaluated with the criteria: 1) what type of the platform is (media sharing, microblogging, social networking) 2) what kinds of content (image, text, video) is posted in the platforms and 3) what kind of attributes are there (Table 1). 4) It is also important to look the number of users to know the popularity of the platforms which helps to know how highly is used the platform. 5) Beside this accessibility of data is an important thing in the assessment of social media platforms to be able to get needed data (Table 1). Earlier studies have used social media platforms like Flickr, Instagram, Twitter and Panoramio also other platforms like TripAdvisor, Geocoaching, and OpenStreetMap were used. Facebook was not used even if they have described that there is massive amount of data (special permission is needed to access data).

The type of content that people post on social media sites varies. Microblogs like Twitter mainly consist of short text content with embedded images and links to other online content, while media-sharing site like Flickr is rich in visual content and related text explanations and comments. A social media post's information content can be broken down into several components: user profile, date/time stamp, geographic information (geotag), certain users' comments and likes and content (image, text, video). In most social media platforms, an Application Programming Interface (API) is used to collect publicly available contents published by the platform's users. The date of the content produced or posted on social media can be included in a social media post also can include information about what users find. The posts can also contain the location where they were taken or posted from and why at that location. Content of the post may also reveal who the users are e.g., gender, age, country of origin. Overall social media content can also be used to answer questions like what? and why? the users posted.

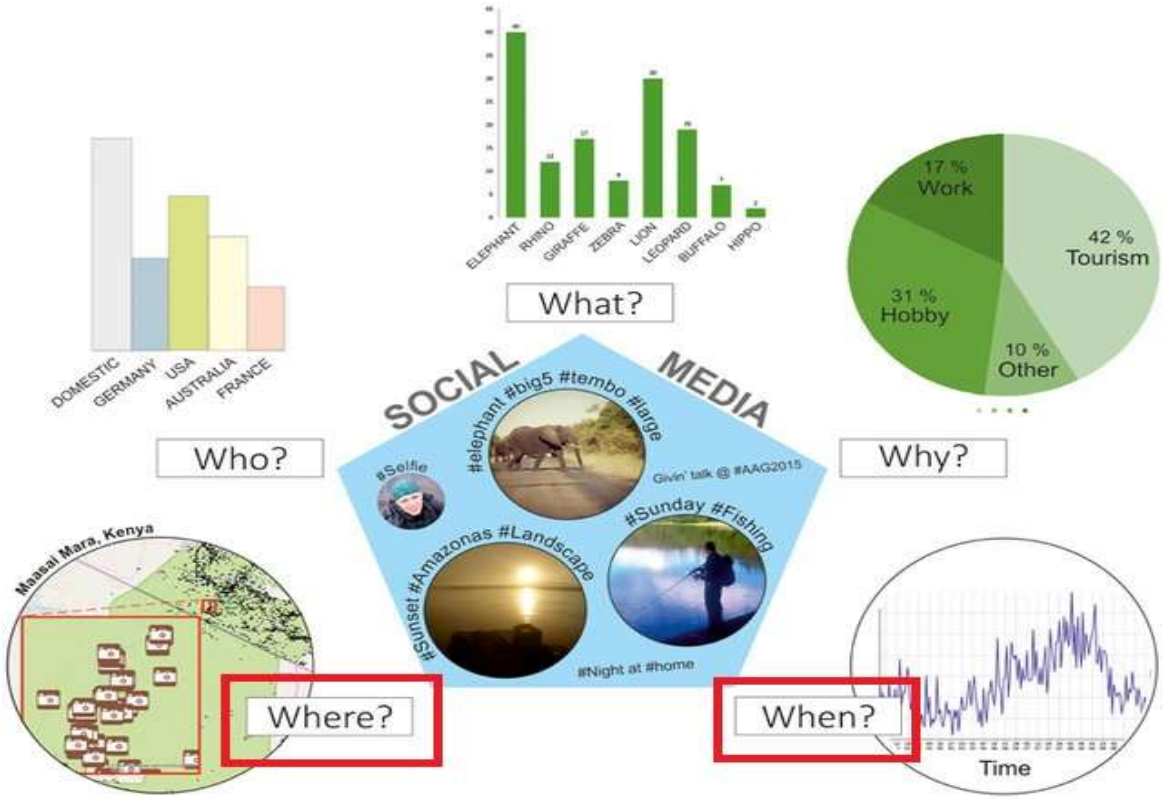

Figure 4.Conceptual framework for the use of social media data (source: (di Minin et al., 2015))

Table 1. Description of selected social media platforms updated from, di Minin et al., (2015); McCay-Peet, L., and Quan-Haase, (2017) and platform statistics and API information were obtained from the platform websites.

Platform	Type of social media	Description	Number of users	API	Data accessibility	Available search parameters used to restrict the content that is downloaded
Flickr	Media sharing platform	-Sharing images and short videos -Popular in professional and nature photographer	Over 60 million monthly users (Last Updated on 2/27/21)	https://www.flickr.com/services/api/	-API is accessible for non-commercial use -Users may attach a license to their photos	Date/time stamp, location, user, keywords, tags
Twitter	Microblogging	-A microblogging social media platform used for sharing short messages (max 280 characters)	353 million monthly active users (as of the first quarter of 2019)	Different API available: https://developer.twitter.com/en/docs/twitter-api	-Twitter open its full tweet archive to academic research product track -Twitter provides a commercial option (Premium and Enterprise search) -Twitter streaming API returns a sample of real-time tweets based on search criteria.	Location, user, keywords, tags, date/time stamp, followers
Instagram	Media sharing	-Photo and small video sharing platform	1 billion monthly active users (as of January 2020)	https://www.instagram.com/developer/	-Have implemented a change in its API, which is basic API permission was disabled on June 29, 2020	User, location, keywords, tags
Facebook	Social networking	-General-purpose social networking -Content: text, image, video	2.7 billion active users monthly (as of the first quarter of 2021)	https://developers.facebook.com/docs/graph-api/	-For the reading of data using the Graph API, special permissions are necessary.	User, page, event, group, place

3.2.3. Social media platform comparison and selection

Researchers can collect data from social media platforms that include time and location for their study using an API. Based on information obtained from the platform websites data from Instagram can be collected but Instagram Legacy API permission ("Basic Permission") was disabled on June 29, 2020, and third-party applications have no longer access to the API. Also, the world's most common website, Facebook, cannot be used in this research due to restricted data access. Whereas, using Flickr API, it is possible to download geotagged and timestamped photos and use for research. APIs can also be used to collect data from Twitter, where millions of individuals openly exchange short text messages. But for Twitter and Flickr alike, users can delete content, and Twitter and Flickr can delete user accounts, so one might not get everything. From the literature review results Flickr take the first place which is used in 42 papers, Twitter takes the second place with 8 papers and Instagram the third with 7 papers (see appendix). In this study Flickr and Twitter platforms are selected as they provide open access to data.

Flickr which was founded in 2004 is one of the oldest social media sites, with a focus on photo sharing, used to store and organize photographs. The Flickr database contains both actual images and picture metadata. This information includes camera settings as well as the date and time of the photo's capture and location attributes is present in the pictures. Table 2 gives an overview and description of the metadata attributes and date taken, latitude and longitude are the main attributes that are used in the study.

Table 2. Flickr metadata selected attributes for this research and their description (Source: Flickr)

Field value	Description	Usage in this research
id	The photo's unique identifier.	Used to uniquely identify the photo
date_taken	The date on which the photo was taken.	Used to know when the picture is taken
accuracy	The level of accuracy of the location information that was recorded. The current range is 1 to 16: World level is 1, Country is ~3, Region is ~6, City is ~11, Street is ~16	Used to understand the accuracy of the location information
url	URL of the photo with different size	Used to check the photos
longtiude	A valid longitude coordinate of the photo location	Used to know the location of the photo
latitude	A valid latitude coordinate of the photo location	Used to know the location of the photo

Twitter is a real-time, highly interactive microblogging service that enables users to post short status updates, known as tweets, that are displayed on timelines. Tweets can include one or more entities in their 280-character text, as well as references to one or more places in the real world. Before using twitter API, it is important to understand the data attributes of Twitter and Table 3 give a description for that. Attributes that have location and date information are the main attributes that are used in this study.

Table 3. Twitter selected data attributes for this research and their description, Twitter API v2 (Source: Twitter)

Field value	Description	Usage in this research
Id	The requested tweet's unique identifier	This can be used to retrieve a particular tweet.
text	The tweet's actual UTF-8 text.	Extraction of keywords
author_id	The user who posted this tweet's unique identifier.	Dataset exchange for peer review
created_at	The tweet's creation time	This field can be used to figure out when a tweet was made, as well as for time-series analysis and other purposes.
geo	If the user mentioned a location in this tweet, this field contains information about that location.	Determine if a tweet is associated with a specified location and its associated geo coordinates.
entities__urls	Expanded media content's	Can be used to see media

3.3. Data Collection and Cleaning

3.3.1. Flickr data collection and cleaning

Flickr is used as a source of information for downloading images taken at the Lake Tana Biosphere Reserve. Data is collected from the Flickr API using R code within the boundaries of the Lake Tana Biosphere Reserve (see Supplementary material). To set the area of interest the shapefile of the Lake Tana Biosphere Reserve was loaded using the 'readOGR' function. The start and the end date of the search were set using the 'photo_search' function by giving the 'mindate_taken' and 'maxdate_taken'. From Flickr 1165 pictures taken in the Lake Tana Biosphere Reserve between 2016-2020 were collected and five years data were preferred to ensure sufficient sample to make robust conclusion. In the collected Flickr data, some images are selfies (example in Figure 5) or are not related to nature visitation. To remove these images, this study uses Kutools for Excel's Insert Pictures from path (URL) to convert the URLs to images. The images were assessed manually, and selfies and off-topic images removed, leaving 978 pictures for the following step. ArcGIS was used to keep one photo point per raster cell (20 m resolution). After that 671 pictures location including urban areas were remained. As the study is on nature visitation ArcGIS was used to remove built up areas using 'Extract Values to Points' tool and 498 pictures location remained for further analysis.

(a) (b)

Figure 5. Examples of selfie pictures: (a) Selfie pictures that are kept because it is trying to show people enjoying nature. (b) This kind of selfie pictures are removed because the intension is not to show nature

3.3.2. Twitter data collection and cleaning

Twitter has introduced its new Academic Research Product Track v2. API endpoint on January 26/2021. This version allows researchers to access large volumes of twitter data over a long-time range. To make any request to the Twitter API a developer application is needed and an application for gathering tweets must be created. After the application has been created and approved Twitter provides for the user API keys and bearer token and with the bearer token, the scrapping process can be performed. In this study the data collection process is performed by using the 'academictwitteR' package obtained from (Barrie & Ho, 2021) using R code (see Supplementary material). Twitter allows to query tweets originating from within a particular geographical buffer. In this study the longitude and latitude of the southwest and then the northeast corners of the bounding box of the Lake Tana Biosphere Reserve were specified and collect all tweets within the bounding box. In the 'start_tweets' and 'end_tweets' the start and end date of the search were set. The study area was divided in several parts because width and height of the bounding box must be less than 25miles to collect tweets and a website [https://boundingbox.klokantech.com/] was used to find the bounding box. From Twitter 545 tweets in the Lake Tana Biosphere Reserve between 2016-2020 were collected. ArcGIS was used to keep one tweet per raster cell (20 m resolution). As the study is on nature visitation tweets in built up areas are removed using 'Extract Values to Points' tool in ArcGIS, and 200 tweets location remained for further analysis.

3.4. Spatial distribution of Photo User Days and Tweet User Days

It is found that users upload many images on a single visit that is taken on the same location. So, the spatial visitation pattern of the Lake Tana Biosphere Reserve were identified by quantifying the proxy Photo User Day (PUD), which is the number of unique users who upload at least one photo in a specific location on a specific day (Wood et al., 2013). For example, if a user takes a picture at a given location twice, only one need to be considered in the analysis, one photo per user per day in a defined site. The PUD should not be taken as a reliable indicator of actual visitor numbers (Wood et al., 2013). Rather, it should be considered as a proxy for the variation in visitation density for each area, given it does not reflect all previous visitors to the location. The PUD distribution were assessed using a statistical summary and plotting the PUD. Since the PUDs are distributed anywhere in the study area the PUD was calculated with 1 × 1 km grid cell size to know which location has more or less PUD. The spatial visitation pattern was also identified by quantifying the proxy Tweet User Day (TUD), which is the number of unique users who tweets at least one tweet in a specific location on a specific day. The same steps as for PUD were taken to map the TUD distribution.

3.4.1. Location characteristics nature visitation

In order to explain the spatial pattern of PUD and TUD, multiple regression is applied for this study involving several social and biophysical parameters (Table 4). Considering multiple parameters plays a significant role to know the most important parameters that can explain the spatial pattern of the location visited in the study area. The distances were calculated using Euclidean distance tool with 1000m cell size for independent variables. Intersect tool was used to know the land cover type in the tiles and obtain in which land cover type is the PUD based on the majority land cover. Sample data having 50 observations with PUD and 50 without PUD was used in the model to check both with and without PUD. The sample data are in different distance, and this will help to avoid considering PUD that are close to each other. The R function vif() Variance Inflation Factor in [R car package] was used to detect and avoid multicollinearity between variables, and the variables that make a best model are selected using the stepwise AIC(Akaike Information Criteria) approach. The same steps as for PUD were taken to TUD having a sample data, 20 observations with TUD and 20 observations without TUD.

Table 4. Justification/Hypothesized link to visitation for independent variables selection

Independent Variables (km^2)	Justification/Hypothesized link to visitation	Data Source
Distance to road	Overland travel to the biosphere reserve has an opportunity to stop and enjoy the numerous beauties along the road and Bahir Dar and the Blue Nile Falls are connected by many buses each day and most people take these buses to go to the Blue Nile Falls.	Road, boat service, religious site, resort, ferry route, and river data from Open Street Map and Tegegne Sitotaw (PhD candidate at ITC), Land cover type from ESA Climate Change Initiative - Land Cover project 2017
Land cover type	The Biosphere Reserve has different land cover type and is rich in biological diversity and a delight for nature lovers to visit	
Distance to ferry route	Ferries provide an opportunity to combine travel with a lake trip and enjoy the nature.	
Distance to boat service	Lake Tana Marin Transport Authority offers a daily boat service to the Lake Tana, churches, and monasteries and also boats can be easily rented from many places in Bahir Dar to travel with a lake trip and people take the boat to enjoy the lake trip.	
Distance to religious site	There are many religious sites in the Lake Tana Biosphere Reserve and one can overlook the old monasteries and countless religious artifacts in the Biosphere Reserve.	
Distance to resort	The resorts that have good views of the lake and is within easy walking distance of the city centre.	
Distance to river	Natural resources such as various rivers and waterfalls found in the Biosphere Reserve	

3.5. Temporal distribution of Photo User Days and Tweet User Days

In which month people are going mostly to the Lake Tana Biosphere Reserve will be answered in this section. From Flickr and Twitter data there is a field value called date taken and created at that describe on which hour, day, month, and year is the photo taken and the tweet created. This study assessed on which month are the Photo User Day taken and the Tweet User Day created over the year. Temporal distribution of PUD and TUD for months of the year were analysed with the amount of PUD and TUD for the 5 years.

4. RESULTS

4.1. Spatial distribution of Photo User Days and Tweet User Days

From the spatial distribution of the PUD, it is identified that most 1km cells have 2 PUDs and a maximum of 33 PUDs. Most PUDs are clustered in the south (around Bahir Dar) and northeast part of the area and with some PUDs in the Lake Tana. The PUD is in Bahir Dar are based on no-urban observations. The main bus station in Bahir Dar is located in the heart of the city and there are several buses a day that connect Bahir Dar with the Blue Nile Falls. Most PUDs are close to the road and on the way of Blue Nile Falls (Tis-Isat Falls). The PUD per land cover type was identified and at least there was 3 PUD in Bare Areas and a maximum of 237 PUD in Cropland (Table 5). From the spatial distribution of the TUD, it is identified that most 1km cells have 1 TUDs and a maximum of 29 TUDs. Also, the TUD per land cover type was identified and at least there was 13 TUD in Tree cover areas and a maximum of 98 TUD in Cropland. The TUDs on Bahir Dar are based on no-urban observations. Just like the PUD most of TUD are close to the road. Also, there are TUD that are concentrated on the Lake Tana. The spatial distribution of both PUD and TUD are more concentrated around Bahir Dar city, even while observations in built up areas were removed.

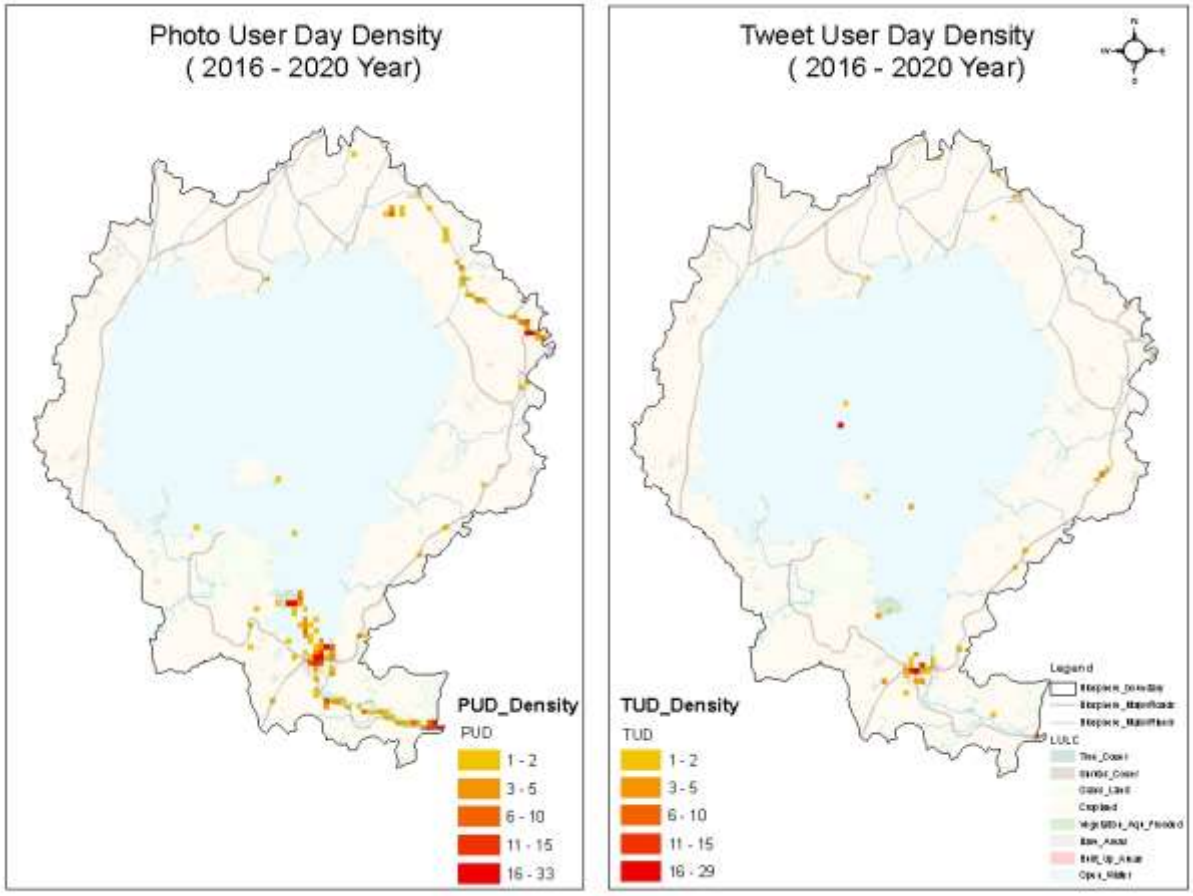

Figure 6. Photo User Day and Tweet User Day Density in the Biosphere Reserve

Table 5. PUD and TUD per land cover type

Land Cover Type	PUD	TUD
Tree cover areas	63	13
Shrubs cover areas	7	-
Grass land	57	35
Cropland	237	98
Vegetation aquatic or flooded	7	-
Bare areas	3	-
Open water	124	54

4.1.1. Location Characteristics nature visitation

The output of the multiple regression model that has two independent variables (land cover type and distance to road) used to explain the response variable PUD_Density. Also, two independent variables (distance to ferry route and distance to road) used to explain the response variable TUD_Density. Tree cover, open water, and distance to road were the significant variables ($p < 0.05$) for PUD_Density in the other hand distance to road and distance to ferry route were the significant variables ($p<0.05$) for the TUD_Density. This implies that the variables are important to explain the spatial pattern of visited locations by the people over the study area. In the PUD model distance to ferry route, distance to boat service, distance to religious site, distance to resort, and distance to river were removed this is because based on the AIC value the amount of information loss by removing the variables is minimum. For TUD model the removed variables are different. This implies that there is a difference in variables to explain the spatial pattern of visited locations between PUD model and TUD model. Also, there is a negative correlation between PUD_Density and distance to road but positive correlation between TUD_Density and distance to road.

From a statistical point of view, using a stepwise approach that is based on AIC in the multiple regression model help to identify the most important independent (explanatory) variables in the models that can explain the intended response (outcome) variable. In this case, the PUD model explains about 21% of PUD_Density and TUD model explains about 31% of TUD_Density (Table 6).

Table 6. Multiple regression model result (LC_Type: Land cover type, Dis_Road: Distance to road, LC_TypeOW: Open water, LC_TypeTC: Tree cover, Dis_ferry: Distance to ferry route)

PUD model		
lm (formula = PUD_Density ~ LC_Type + Dis_Road, data = PUD_model)		
Coefficients:		
	Estimate	Pr(>|t|)
(Intercept)	2.4505	**0.00153**
LC_TypeOW	4.2385	**0.00825**
LC_TypeTC	13.1631	**<0.001**
Dis_Road	-0.4581	**0.00180**
Multiple R-squared: 0.2144		
Adjusted R-squared: 0.1791		

TUD model		
lm (formula = TUD_Density ~ Dis_Road + Dis_ferry, data = TUD_model)		
Coefficients:		
	Estimate	Pr(>|t|)
(Intercept)	4.54899	**0.01317**
Dis_ferry	-0.20439	**0.00766**
Dis_road	0.32474	**0.04692**
Multiple R-squared: 0.3153		
Adjusted R-squared: 0.2646		

4.2. Temporal Distribution of Photo User Days and Tweet User Days

Having the PUD and TUD figure 7, 8 and 9 compares the absolute and relative number of PUD and TUD per month. The charts are based on the timestamps of PUD and TUD which are automatically assigned to the photos when they are taken and to the tweets when they are created. It shows that the monthly distribution of PUD is different. Most of the PUD are from January with 112 PUD and the least PUD from August with 1 PUD. The TUD monthly distribution is nearly the same with 28 TUD in December and 7 TUD in September, October, and November. On the Lake Tana Biosphere Reserve website (https://laketana-biosphere.com/) they have mentioned festivals are celebrated in the biosphere. The most vibrant festivals in Ethiopia are Timket (Ethiopian Epiphany) on January 19, Ethiopian Christmas on January 7, Meskel (finding of the True Cross) on the September 27 and Ethiopian New Year celebrations

on September 11. The January month, that has two colourful festivals, has a higher number in both PUD and TUD. Early October is the best time to visit Lake Tana since the lake is at its fullest, and the Tis Isat Falls are at their most magnificent but there is less PUD and TUD in October (Figure 11). People go for bird watching in the Ethiopian winter (December to February) since resident species are supplemented with large numbers of migrating birds (When to Visit Lake Tana Biosphere Reserve - Lake Tana Biosphere Reserve, n.d.). Because the sun shines almost every day of the year in Ethiopia, even during the rainy season, Lake Tana can be visited all year.

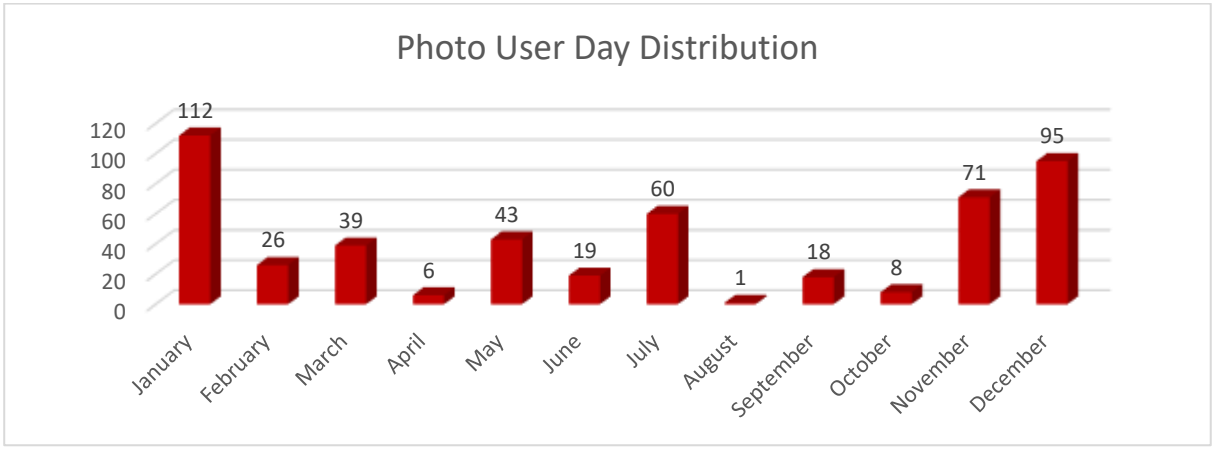

Figure 7. PUD Distribution by month

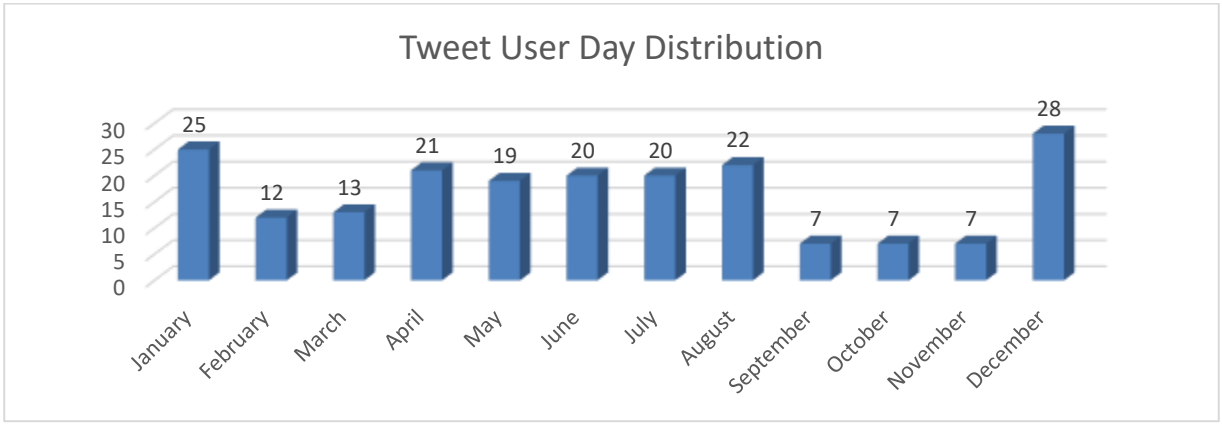

Figure 8. TUD Distribution by month

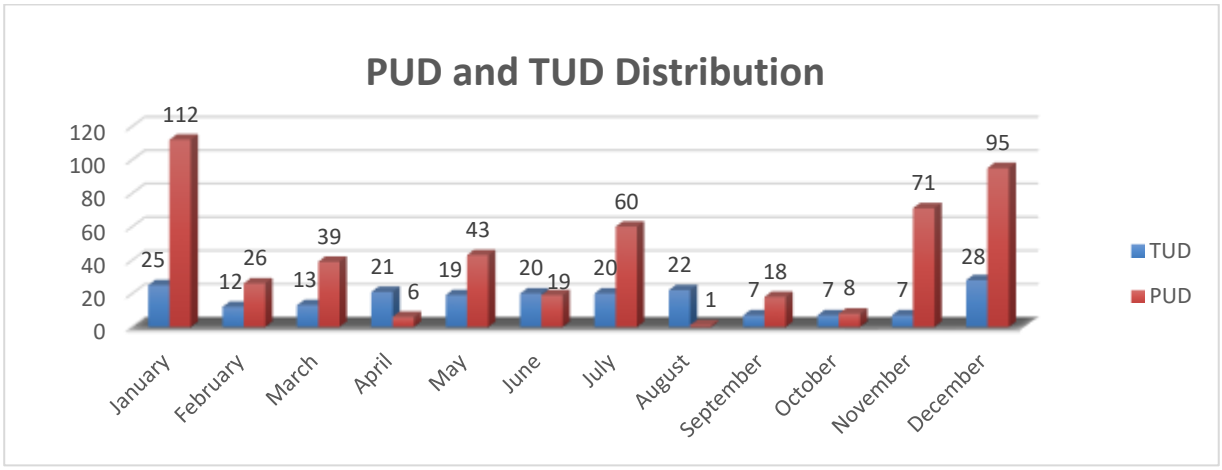

Figure 9. PUD and TUD distribution by month

5. DISCUSSION

5.1. Social media data for Cultural Ecosystem Services

CES are difficult to quantify however, in this study it is showed that CES can be used as proxy to spatially and temporally quantified using social media data. Information available through crowdsourced data in the form of pictures, location and textual data reflects how individuals are appreciating their natural environment. The use of social media data to evaluate nature visitation is yet new research. While this area is still new, several authors, journals, places, and platforms are gaining competence in using different methods to analyse this type of data. Compared to the traditional methods like direct observations, surveys and interviews social media data gives a wider range of spatial and temporal data to assess nature visitation (Teles da Mota, V.; Pickering, 2018). However, such kind of data is more apparent in natural areas with important tourist attractions or areas of high human activity like roads than locations with minor tourist attractions. This was evident in this study with PUD concentrated along road network. In citizen science portal, observation concentrations may indicate a biodiversity sampling bias but they can be interpreted as good evidence for a large CES supply in a cultural sense (Havinga et al., 2020).

Spatial accuracy like accuracy of geotag can be influenced by mobile coverage, recording device, GPS visibility, and manual geotagging procedures (Li et al., 2013). Location associated with tweets and photos maybe automatically captured from built in GPS receivers in mobile devices or can be manually filled by a user (Li et al., 2013). In this study the locations for both tweets and photos are resolved by keeping one photo/tweet point per raster cell (20 m resolution) but it should be expected that the accuracy of the locations is dependent on the accuracy of GPS or the map scale when the user specifies the locations. Using the location of photos and spatial data it is possible to know which sites people visit. In this case people might take photos from far away and what is valued is unknown without doing image interpretation, but they still visited that place because of the views/value of the area. Social media data sampled users may be biased in age, gender, socioeconomic status, and education. Younger, wealthier, and more educated users dominate the most popular platforms (Aaron Smith et al., 2018) and this was the limitation of this study not identifying who the users are. Deep learning algorithms may be used to analyse user profiles in order to compensate for these limitations (Hausmann et al., 2017). Despite the limitation this data source can be used almost anywhere including developing and developed countries, as well as data poor and data rich locations to indicate visitation (Wood et al., 2020).

Privacy issues and ethical use must be considered while using social media data and personal information. Users on social media platforms have a variety of privacy settings, including the option to "opt-in" to geotagging (Havinga et al., 2020). Researchers must, however, assess whether technology providers have provided users with enough understanding and control over their data. Users' sharing of information by social media platforms may be governed by a variety of policies. For example, some encourage users to disclose their location information. Furthermore, because services may be terminated, there is no guarantee that the data will be available in the future (for example, Panoramio was shut in November 2016) and API conditions may change (for example, the Instagram API was changed which is basic API permission was disabled on June 29, 2020).

5.2. Nature Visitation in Lake Tana Biosphere Reserve

The spatial and temporal pattern of nature visitation in Lake Tana Biosphere Reserve was assessed using the PUD and TUD concept. The PUD and TUD will help to control the analysis not to be biased towards active users. Using the location of photos and tweets with combination of other geospatial data it is possible

to know which sites people visit. Mapping the locations of PUD and TUD can provide information into areas that have larger number of photos and tweets. Knowing location that have a larger number of photos and tweets is important since it represents an area that may require monitoring and the building of supporting infrastructure. The results from this study show that most of PUD and TUD are close to the road and in the regression analysis accessibility to the road has an influence on the PUD and TUD density. Also, from the land cover type tree cover areas have an influence on PUD density in the regression analysis. These findings can be used with other data sources and on-site monitoring to determine if these locations have more visitor, if so, what can be done about it in tourism planning and decision making. In the study it is found that the R-Squared value in TUD model is greater than the PUD model. Beside this the adjusted R-Squared has decrease in both models. R-Squared increases even when you add variables which are not related to the dependent variable. But adjusted R-Squared take care of that as it decreases whenever you add variables that are not related to the dependent variable, thus after taking care it is likely to decrease (Leach et al., 2007). So, relating to this study finding variables that can explain the spatial pattern better is needed. PUD and TUD was used to detect the most popular month for visitation in the Lake Tana Biosphere Reserve. It is found that the area is highly visited in month where festivals are celebrated. Knowing the temporal pattern will help to increase facilities on the area that are important for visitation on the highly visitation months. Also, it will help conservation authorities in managing areas with high concentration of visitation. Overall using social media data to assess nature visitation is new for this study area so, there is a need of more research and practical development before using social media data for monitoring visitation.

5.3. Recommendations

Social media platforms change rapidly in a way to access, collect and analyse the data so it is recommended to check the platforms if there is any change. In addition, further research is needed to determine the impact of differences in who posts what to which social media platforms and why. API access on social media sites can change on a frequent basis. Twitter and Instagram, for example, have both modified their access levels in recent years so it is recommended to check API access. Uncertainties in crowdsourced data's location accuracy must also be considered. It would be good to have studies that compare social media data with other sources in the future in the Lake Tana Biosphere Reserve to further understand how social media data may complement and build on existing sources of information. This study was focused on the spatial and temporal pattern so, combining this pattern, like dividing the spatial pattern in patterns per month, could give valuable results. Using social media data to assess nature visitation in the Lake Tana Biosphere reserve is a new research so, automated procedures, such as image content analysis and processing vast data sources, needs to be investigated.

6. CONCLUSION

The objective of this research was to assess nature visitation on the Lake Tana Biosphere Reserve using social media data. To reach the objective, a research questions What are the key characteristics and differences between Flickr and Twitter with respect to structure and content? What are the spatial and temporal pattern of nature visitation in the Lake Tana Biosphere Reserve? and similarities and differences in spatial and temporal pattern between the platforms were addressed. Research using social media have experienced difficulties, such as fluctuating platform popularity, data access, and answering the research questions for certain regions, activities, and values. However, as social media becomes more prevalent in our lives, it may become an increasingly important source of data for monitoring nature visitation. For this study, 1165 photos taken in the study area between 2016-2020 were collected from Flickr,

and 498 photo locations were analysed. The results presented in this research suggest that social media data may act as an additional information source for planning as well as managing the Lake Tana Biosphere Reserve and understanding nature visitation. Lake Tana Biosphere Reserve's spatial distribution of visitation were assessed by PUD and TUD density. Also, with the multiple regression analysis using several parameters helps to know which parameters have an influence on the PUD and TUD density. Temporal distribution of PUD and TUD for months of the year were extracted and analysed the amount of PUD and TUD for months. There was a high number of PUD and TUD in months with festivals that is why people visited mostly also people go for bird watching in the Ethiopian winter (December to February). The results could, for example, assist tourism managers and planners in identifying popular tourist destinations and those with the potential to welcome more visitors. Conservation authorities could better understand spatiotemporal patterns in tourist preferences by continuously monitoring social media data combining with other geospatial data.

APPENDIX: LITERATURE REVIEW

Publications that used social media platforms to assess nature visitation

Publications	Flickr (N=42)	Twitter (N=7)	Instagram (N=8)	Panoramio (N=8)
(Allan et al., 2015; Barros et al., 2020; Fisher et al., 2018; Levin et al., 2017; Levin et al., 2015; Long et al., 2021; Martinez-Harms et al., 2018; Moreno-Llorca et al., 2020; Richards & Friess, 2015; Sessions et al., 2016; Sinclair et al., 2018, 2020; Sonter et al., 2016; Spalding et al., 2017; Thiagarajah et al., 2015; Wartmann et al., 2018; Willemen et al., 2015; Wood et al., 2013; Yoshimura & Hiura, 2017; Cunha et al., 2018; Dunkel, 2015; Ghermandi, 2018; Hale, 2018; Keeler et al., 2015; Lee et al., 2014; Richards et al., 2018; Tenerelli et al., 2016, 2017; Walden-Schreiner, Leung, et al., 2018; Walden-Schreiner, Rossi, et al., 2018)	✓			
(Ghermandi, 2016; Oteros-Rozas et al., 2018; Tieskens et al., 2018; Figueroa-Alfaro & Tang, 2017)	✓			✓
(Ghermandi et al., 2020)	✓	✓		✓
(Hamstead et al., 2018)	✓	✓		
(Van Zanten et al., 2016; (Angradi et al., 2018)	✓		✓	✓
(Hausmann et al., 2018)	✓		✓	
(Tenkanen et al., 2017) Wood et al., 2020)	✓	✓	✓	
(Teles da Mota & Pickering, 2020)	✓	✓	✓	✓
(Conti & Lexhagen, 2020) Hausmann et al., 2017)			✓	
(Becken et al., 2017, 2020)		✓		

LIST OF REFERENCES

Aaron Smith, B., Anderson, M., Smith, A., Caiazza, T., & Manager, C. (2018). *Social Media Use in 2018* (Vol. 1). Pew Research Center. www.pewresearch.org.

Alemayu Wassie Eshete. (2007). Ethiopian Church Forests Opportunities and Challenges for Restauration. In *(Doctoral dissertation, Wageningen University, Wageningen, The Netherlands)*.

Allan, J. D., Smith, S. D. P., McIntyre, P. B., Joseph, C. A., Dickinson, C. E., Marino, A. L., Biel, R. G., Olson, J. C., Doran, P. J., Rutherford, E. S., Adkins, J. E., & Adeyemo, A. O. (2015). Using cultural ecosystem services to inform restoration priorities in the Laurentian Great Lakes. *Frontiers in Ecology and the Environment*, *13*(8), 418–424. https://doi.org/10.1890/140328

Angradi, T. R., Launspach, J. J., & Debbout, R. (2018). Determining preferences for ecosystem benefits in Great Lakes Areas of Concern from photographs posted to social media. *Journal of Great Lakes Research*, *44*(2), 340–351. https://doi.org/10.1016/j.jglr.2017.12.007

Barrie, C., & Ho, J. (2021). academictwitteR: an R package to access the Twitter Academic Research Product Track v2 API endpoint. *Journal of Open Source Software*, *6*(62), 3272. https://doi.org/10.21105/joss.03272

Barros, C., Moya-Gómez, B., & Gutiérrez, J. (2020). Using geotagged photographs and GPS tracks from social networks to analyse visitor behaviour in national parks. *Current Issues in Tourism*, *23*(10), 1291–1310. https://doi.org/10.1080/13683500.2019.1619674

Becken, S., Alaei, A. R., & Wang, Y. (2020). Benefits and pitfalls of using tweets to assess destination sentiment. *Journal of Hospitality and Tourism Technology*, *11*(1), 19–34. https://doi.org/10.1108/JHTT-09-2017-0090

Becken, S., Stantic, B., Chen, J., Alaei, A. R., & Connolly, R. M. (2017). Monitoring the environment and human sentiment on the Great Barrier Reef: Assessing the potential of collective sensing. *Journal of Environmental Management*, *203*, 87–97. https://doi.org/10.1016/j.jenvman.2017.07.007

Bires, Z., & Raj, S. (2019). Determinants of environmental conservation in Lake Tana Biosphere Reserve, Ethiopia. *Heliyon*. https://doi.org/10.1016/j.heliyon.2019.e01997

Canedoli, C., Bullock, C., Collier, M. J., Joyce, D., & Padoa-Schioppa, E. (2017). public participatory mapping of cultural ecosystem services: Citizen perception and park management in the Parco Nord of Milan (Italy). *Sustainability (Switzerland)*, *9*(6). https://doi.org/10.3390/su9060891

Codato, D., Pappalardo, S. E., & de Marchi, M. (2017). Participatory GIS in Mapping Ecosystem Services: Two Case Studies from High-Biodiversity Regions in Italy and Peru. *GI_Forum*, *1*(January), 78–96. https://doi.org/10.1553/giscience2017_02_s78

Conti, E., & Lexhagen, M. (2020). Instagramming nature-based tourism experiences: a netnographic study of online photography and value creation. *Tourism Management Perspectives*, *34*, 100650. https://doi.org/10.1016/j.tmp.2020.100650

Crossman, N. D., Burkhard, B., Nedkov, S., Willemen, L., Petz, K., Palomo, I., Drakou, E. G., Martín-Lopez, B., McPhearson, T., Boyanova, K., Alkemade, R., Egoh, B., Dunbar, M. B., & Maes, J. (2013). A blueprint for mapping and modelling ecosystem services. *Ecosystem Services*, *4*, 4–14. https://doi.org/10.1016/j.ecoser.2013.02.001

Cunha, J., Elliott, M., & Ramos, S. (2018). Linking modelling and empirical data to assess recreation services provided by coastal habitats: The case of NW Portugal. *Ocean and Coastal Management*, *162*, 60–70. https://doi.org/10.1016/j.ocecoaman.2017.12.022

Delgado-Aguilar, M. J., Konold, W., & Schmitt, C. B. (2017). Community mapping of ecosystem services in tropical rainforest of Ecuador. *Ecological Indicators*, *73*, 460–471. https://doi.org/10.1016/j.ecolind.2016.10.020

di Minin, E. Di, Tenkanen, H., & Toivonen, T. (2015). Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, *3*(SEP). https://doi.org/10.3389/fenvs.2015.00063

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, *142*, 173–186. https://doi.org/10.1016/j.landurbplan.2015.02.022

Figueroa-Alfaro, R. W., & Tang, Z. (2017). Evaluating the aesthetic value of cultural ecosystem services by mapping geo-tagged photographs from social media data on Panoramio and Flickr. *Journal of Environmental Planning and Management*, *60*(2), 266–281.

https://doi.org/10.1080/09640568.2016.1151772
Fisher, D. M., Wood, S. A., White, E. M., Blahna, D. J., Lange, S., Weinberg, A., Tomco, M., & Lia, E. (2018). Recreational use in dispersed public lands measured using social media data and on-site counts. *Journal of Environmental Management*, *222*, 465–474. https://doi.org/10.1016/j.jenvman.2018.05.045
Getnet Fetene, B. D. P. O. (2017). *Reference Manual Lake Tana Biosphere Reserve Day*.
Ghermandi, A. (2016). Analysis of intensity and spatial patterns of public use in natural treatment systems using geotagged photos from social media. *Water Research*, *105*, 297–304. https://doi.org/10.1016/j.watres.2016.09.009
Ghermandi, A. (2018). Integrating social media analysis and revealed preference methods to value the recreation services of ecologically engineered wetlands. *Ecosystem Services*, *31*, 351–357. https://doi.org/10.1016/j.ecoser.2017.12.012
Ghermandi, A., Sinclair, M., Fichtman, E., & Gish, M. (2020). Novel insights on intensity and typology of direct human-nature interactions in protected areas through passive crowdsourcing. *Global Environmental Change*, *65*, 102189. https://doi.org/10.1016/j.gloenvcha.2020.102189
Hale, B. W. (2018). Mapping Potential Environmental Impacts from Tourists Using Data from Social Media: A Case Study in the Westfjords of Iceland. *Environmental Management*, *62*(3), 446–457. https://doi.org/10.1007/s00267-018-1056-z
Hamstead, Z. A., Fisher, D., Ilieva, R. T., Wood, S. A., McPhearson, T., & Kremer, P. (2018). Geolocated social media as a rapid indicator of park visitation and equitable park access. *Computers, Environment and Urban Systems*, *72*, 38–50. https://doi.org/10.1016/j.compenvurbsys.2018.01.007
Hausmann, A., Toivonen, T., Heikinheimo, V., Tenkanen, H., Slotow, R., & Minin, E. DI. (2017). Social media reveal that charismatic species are not the main attractor of ecotourists to sub-Saharan protected areas. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-00858-6
Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., & Di Minin, E. (2018). Social Media Data Can Be Used to Understand Tourists' Preferences for Nature-Based Experiences in Protected Areas. In *Conservation Letters* (Vol. 11, Issue 1, p. e12343). Wiley-Blackwell. https://doi.org/10.1111/conl.12343
Havinga, I., Bogaart, P. W., Hein, L., & Tuia, D. (2020). Defining and spatially modelling cultural ecosystem services using crowdsourced data. *Ecosystem Services*, *43*(April), 101091. https://doi.org/10.1016/j.ecoser.2020.101091
Helka, J. (2016). *The importance of cultural ecosystem services and their evaluation in the context of ecosystem service assessment frameworks Comparative study of users' perceptions (preferences) of cultural ecosystem services and disservices in and outside the city of Le*. *April*.
Hirons, M., Comberti, C., & Dunford, R. (2016). Valuing Cultural Ecosystem Services. *Annual Review of Environment and Resources*, *41*, 545–574. https://doi.org/10.1146/annurev-environ-110615-085831
Kalmbach, E. (2017). *Establishment of the Lake Tana Biosphere Reserve within the UNESCO World Network of Biosphere Reserves* (pp. 523–544). Springer, Cham. https://doi.org/10.1007/978-3-319-45755-0_29
Kassahun, T., & Bender, S. (2019). Saving the Last Endemic-Church Forests in Ethiopia: The Case of Lake Tana Biosphere Reserve. In *Climate Change Management* (pp. 195–210). Springer. https://doi.org/10.1007/978-3-319-98681-4_12
Keeler, B. L., Wood, S. A., Polasky, S., Kling, C., Filstrup, C. T., & Downing, J. A. (2015). Recreational demand for clean water: Evidence from geotagged photographs by visitors to lakes. *Frontiers in Ecology and the Environment*, *13*(2), 76–81. https://doi.org/10.1890/140124
Kim, Y., Kim, C. ki, Lee, D. K., Lee, H. woo, & Andrada, R. I. T. (2019). Quantifying nature-based tourism in protected areas in developing countries by using social big data. *Tourism Management*, *72*(February 2018), 249–256. https://doi.org/10.1016/j.tourman.2018.12.005
Leach, L. F., Henson, R. K., Finch, W. H., Fraas, J. W., Newman, I., & Walker, D. A. (2007). *Multiple Linear Regression Viewpoints Volume 33 • Number 1 • Fall*.
Lee, I., Cai, G., & Lee, K. (2014). Exploration of geo-tagged photos through data mining approaches. *Expert Systems with Applications*, *41*(2), 397–405. https://doi.org/10.1016/j.eswa.2013.07.065
Levin, N., Kark, S., & Crandall, D. (2015). Where have all the people gone? Enhancing global conservation using night lights and social media. *Ecological Applications*, *25*(8), 2153–2167. https://doi.org/10.1890/15-0113.1
Levin, N., Lechner, A. M., & Brown, G. (2017). An evaluation of crowdsourced information for assessing

the visitation and perceived importance of protected areas. *Applied Geography*, *79*, 115–126. https://doi.org/10.1016/j.apgeog.2016.12.009
Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of twitter and flickr. *Cartography and Geographic Information Science*. https://doi.org/10.1080/15230406.2013.777139
Long, P. R., Nogué, S., Benz, D., & Willis, K. J. (2021). Devising a method to remotely model and map the distribution of natural landscapes in Europe with the greatest recreational amenity value (cultural services). *Frontiers of Biogeography*, *13*(1), 1–13. https://doi.org/10.21425/F5FBG47737
Martinez-Harms, M. J., Bryan, B. A., Wood, S. A., Fisher, D. M., Law, E., Rhodes, J. R., Dobbs, C., Biggs, D., & Wilson, K. A. (2018). Inequality in access to cultural ecosystem services from protected areas in the Chilean biodiversity hotspot. *Science of the Total Environment*, *636*, 1128–1138. https://doi.org/10.1016/j.scitotenv.2018.04.353
McCay-Peet, L., and Quan-Haase, A. (2017). What is social media and what questions can social media research help us answer? In *The SAGE Handbook of Social Media Research Methods* (pp. 13–26). https://books.google.nl/books?hl=en&lr=&id=9oewDQAAQBAJ&oi=fnd&pg=PA13&ots=eOJPq4sWDN&sig=t6e1o4LOjDOpM-lc0X_Xi2LZTF0&redir_esc=y#v=onepage&q&f=false
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. In *PLoS Medicine* (Vol. 6, Issue 7, p. e1000097). Public Library of Science. https://doi.org/10.1371/journal.pmed.1000097
Moreno-Llorca, R., F. Méndez, P., Ros-Candeira, A., Alcaraz-Segura, D., Santamaría, L., Ramos-Ridao, Á. F., Revilla, E., Bonet-García, F. J., & Vaz, A. S. (2020). Evaluating tourist profiles and nature-based experiences in Biosphere Reserves using Flickr: Matches and mismatches between online social surveys and photo content analysis. *Science of the Total Environment*, *737*, 140067. https://doi.org/10.1016/j.scitotenv.2020.140067
Müller, S. M., Peisker, J., Bieling, C., Linnemann, K., Reidl, K., & Schmieder, K. (2019). The importance of cultural ecosystem services and biodiversity for landscape visitors in the biosphere reserve Swabian Alb (Germany). *Sustainability (Switzerland)*, *11*(9), 1–23. https://doi.org/10.3390/su11092650
Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., & Plieninger, T. (2018). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators*, *94*, 74–86. https://doi.org/10.1016/j.ecolind.2017.02.009
Plieninger, T., Dijks, S., Oteros-Rozas, E., & Bieling, C. (2013). Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, *33*, 118–129. https://doi.org/10.1016/j.landusepol.2012.12.013
Richards, D. R., & Friess, D. A. (2015). A rapid indicator of cultural ecosystem service usage at a fine spatial scale: Content analysis of social media photographs. *Ecological Indicators*, *53*, 187–195. https://doi.org/10.1016/j.ecolind.2015.01.034
Richards, D. R., Tunçer, B., & Tunçer, B. (2018). Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services*, *31*, 318–325. https://doi.org/10.1016/j.ecoser.2017.09.004
Sessions, C., Wood, S. A., Rabotyagov, S., & Fisher, D. M. (2016). Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs. *Journal of Environmental Management*, *183*, 703–711. https://doi.org/10.1016/j.jenvman.2016.09.018
Sinclair, M., Ghermandi, A., & Sheela, A. M. (2018). A crowdsourced valuation of recreational ecosystem services using social media data: An application to a tropical wetland in India. *Science of the Total Environment*, *642*, 356–365. https://doi.org/10.1016/j.scitotenv.2018.06.056
Sinclair, M., Mayer, M., Woltering, M., & Ghermandi, A. (2020). Using social media to estimate visitor provenance and patterns of recreation in Germany's national parks. *Journal of Environmental Management*, *263*, 110418. https://doi.org/10.1016/j.jenvman.2020.110418
Sonter, L. J., Watson, K. B., Wood, S. A., & Ricketts, T. H. (2016). Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. *PLoS ONE*, *11*(9), e0162372. https://doi.org/10.1371/journal.pone.0162372
Spalding, M., Burke, L., Wood, S. A., Ashpole, J., Hutchison, J., & zu Ermgassen, P. (2017). Mapping the

global value and distribution of coral reef tourism. *Marine Policy*, *82*, 104–113. https://doi.org/10.1016/j.marpol.2017.05.014
Teles da Mota, V.; Pickering, C. (2018). *How can we use social media to know more about visitors to natural areas*. *2018*, 72–74.
Teles da Mota, V., & Pickering, C. (2020). Using social media to assess nature-based tourism: Current research and future trends. *Journal of Outdoor Recreation and Tourism*, *30*, 100295. https://doi.org/10.1016/j.jort.2020.100295
Tenerelli, P., Demšar, U., & Luque, S. (2016). Crowdsourcing indicators for cultural ecosystem services: A geographically weighted approach for mountain landscapes. *Ecological Indicators*, *64*, 237–248. https://doi.org/10.1016/j.ecolind.2015.12.042
Tenerelli, P., Püffel, C., & Luque, S. (2017). Spatial assessment of aesthetic services in a complex mountain region: combining visual landscape properties with crowdsourced geographic information. *Landscape Ecology*, *32*(5), 1097–1115. https://doi.org/10.1007/s10980-017-0498-7
Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., & Toivonen, T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-18007-4
Thiagarajah, J., Wong, S. K. M., Richards, D. R., & Friess, D. A. (2015). Historical and contemporary cultural ecosystem service values in the rapidly urbanizing city state of Singapore. *Ambio*, *44*(7), 666–677. https://doi.org/10.1007/s13280-015-0647-7
Tieskens, K. F., Van Zanten, B. T., Schulp, C. J. E., & Verburg, P. H. (2018). Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape. *Landscape and Urban Planning*, *177*, 128–137. https://doi.org/10.1016/j.landurbplan.2018.05.002
Troy, A., & Wilson, M. A. (2006). Mapping ecosystem services: Practical challenges and opportunities in linking GIS and value transfer. *Ecological Economics*, *60*(2), 435–449. https://doi.org/10.1016/j.ecolecon.2006.04.007
Van Zanten, B. T., Van Berkel, D. B., Meentemeyer, R. K., Smith, J. W., Tieskens, K. F., & Verburg, P. H. (2016). Continental-scale quantification of landscape values using social media data. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), 12974–12979. https://doi.org/10.1073/pnas.1614158113
Walden-Schreiner, C., Leung, Y. F., & Tateosian, L. (2018). Digital footprints: Incorporating crowdsourced geographic information for protected area management. *Applied Geography*, *90*, 44–54. https://doi.org/10.1016/j.apgeog.2017.11.004
Walden-Schreiner, C., Rossi, S. D., Barros, A., Pickering, C., & Leung, Y. F. (2018). Using crowd-sourced photos to assess seasonal patterns of visitor use in mountain-protected areas. *Ambio*, *47*(7), 781–793. https://doi.org/10.1007/s13280-018-1020-4
Wartmann, F. M., Acheson, E., & Purves, R. S. (2018). Describing and comparing landscapes using tags, texts, and free lists: an interdisciplinary approach. *International Journal of Geographical Information Science*, *32*(8), 1572–1592. https://doi.org/10.1080/13658816.2018.1445257
When to visit Lake Tana Biosphere Reserve - Lake Tana Biosphere Reserve. (n.d.). Retrieved June 16, 2021, from https://www.laketana-biosphere.com/visit/when-to-visit/
Willemen, L., Cottam, A. J., Drakou, E. G., & Burgess, N. D. (2015). Using social media to measure the contribution of red list species to the nature-based tourism potential of african protected areas. *PLoS ONE*, *10*(6), e0129785. https://doi.org/10.1371/journal.pone.0129785
Wondie, A. (2018). Ecological conditions and ecosystem services of wetlands in the Lake Tana Area, Ethiopia. *Ecohydrology and Hydrobiology*, *18*(2), 231–244. https://doi.org/10.1016/j.ecohyd.2018.02.002
Wood, S. A., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports*, *3*(1), 1–7. https://doi.org/10.1038/srep02976
Wood, S. A., Winder, S. G., Lia, E. H., White, E. M., Crowley, C. S. L., & Milnor, A. A. (2020). Next-generation visitation models using social media to estimate recreation on public lands. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-70829-x
Worku, M. (2017). Lake Tana as Biosphere Reserve: Review. *Journal of Tourism & Hospitality*. https://doi.org/10.4172/2167-0269.1000310
Yoshimura, N., & Hiura, T. (2017). Demand and supply of cultural ecosystem services: Use of geotagged

photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services*, *24*, 68–78. https://doi.org/10.1016/j.ecoser.2017.02.009
Zur Heide, F. (2012). *Feasibility Study for a Lake Tana Biosphere Reserve, Ethiopia*. www.bfn.de