Ecological Indicators 96 (2019) 505–515

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

Mapping cultural ecosystem services 2.0 – Potential and shortcomings from unlabeled crowd sourced images

Heera Lee[a,b,*], Bumsuk Seo[b,c], Thomas Koellner[d], Sven Lautenbach[a,e]

[a] Faculty of Agriculture, Land Use Modeling and Ecosystem Services, University of Bonn, Nussallee 1, D-53115 Bonn, Germany
[b] Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Kreuzeckbahnstr. 19, D-82467 Garmisch-Partenkirchen, Germany
[c] Institute of Environmental Research, Kangwon National University, Gangwondaehak-ro 1, 24341 Chuncheon, Republic of Korea
[d] Professorship of Ecological Services, Bayreuth Center of Ecology and Environmental Research BayCEER, University of Bayreuth, Universitaetsstrasse 30, D-95440 Bayreuth, Germany
[e] GIScience Research Group, Institute of Geography, University of Heidelberg, D-69120 Heidelberg, Germany

ARTICLE INFO

Keywords:
Social media photo services
Cloud computing
Tagging
Keywords network analysis
Mapping ecosystem services
Big data analysis

ABSTRACT

The volume of accessible geotagged crowdsourced photos has increased. Such data include spatial, temporal, and thematic information on recreation and outdoor activities, thus can be used to quantify the demand for cultural ecosystem services (CES). So far photo content has been analyzed based on user-labeled tags or the manual labeling of photos. Both approaches are challenged with respect to consistency and cost-efficiency, especially for large-scale studies with an enormous volume of photos. In this study, we aim at developing a new method to analyze the content of large volumes of photos and to derive indicators of socio-cultural usage of landscapes. The method uses machine-learning and network analysis to identify clusters of photo content that can be used as an indicator of cultural services provided by landscapes. The approach was applied in the Mulde river basin in Saxony, Germany. All public Flickr photos (n = 12,635) belonging to the basin were tagged by deep convolutional neural networks through a cloud computing platform, Clarifai. The machine-predicted tags were analyzed by a network analysis that leads to nine hierarchical clusters. Those clusters were used to distinguish between photos related to CES (65%) and not related to CES (35%). Among the nine clusters, two clusters were related to CES: 'landscape aesthetics' and 'existence'. This step allowed mapping of different aspects of CES and separation of non-relevant photos from further analysis. We further analyzed the impact of protected areas on the spatial pattern of CES and not-related CES photos. The presence of protected areas had a significant positive impact on the areas with both 'landscape aesthetics' and 'existence' photos: the total number of days in each mapping unit where at least one photo was taken by a user ('photo-user-day') increased with the share of protected areas around the location. The presented approach has shown its potential for reliable mapping of socio-cultural uses of landscapes. It is expected to scale well with large numbers of photos and to be easily transferable to different regions.

1. Introduction

Quantification of ecosystem services (ES) delivery is essential for the assessment of trade-offs of land use decisions. Cultural ecosystem services (CES) are the most anthropocentric and subjective ES, which makes them particularly difficult to quantify (Daniel et al., 2012; Milcu et al., 2013; Gliozzo et al., 2016; La Rosa et al., 2016). A number of previous CES studies examined stated preferences based on survey data (Gee and Burkhard, 2010; van Berkel and Verburg, 2014) and interviews (Plieninger et al., 2013). Individual surveys and interviews are advantageous as they encourage participation of the local stakeholders in a CES valuation (von Heland and Folke, 2014; Delgado-Aguilar et al., 2017). Also, participatory mapping such as public participation GIS (PPGIS) enhances public involvement in identifying spatially explicit information on CES provision (Brown and Fagerholm, 2015). Given that only 21% of reviewed CES studies used spatial information (La Rosa et al., 2016), PPGIS provides an important step forward in the use of spatial data for CES. Yet surveys are still often

* Corresponding author at: Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Kreuzeckbahnstr. 19, D-82467 Garmisch-Partenkirchen, Germany.
E-mail address: heera.lee@kit.edu (H. Lee).

https://doi.org/10.1016/j.ecolind.2018.08.035
Received 11 December 2017; Received in revised form 7 August 2018; Accepted 16 August 2018
1470-160X/

expensive to conduct and have a limited scope on time and space (Norton et al., 2012; Wood et al., 2013). Furthermore, they can be biased as stated preferences often do not correspond with revealed preferences (Cord et al., 2015).

Recently an alternative indicator for preferences on landscape aesthetics and recreational activities has been introduced to overcome the limitations of stated preferences measures. Social media databases of geotagged photos that have been uploaded to crowdsourcing photo repositories (e.g., *Flickr* and *Panoramio*) have been used to understand socio-cultural usages of landscapes (Keeler et al., 2015; Gliozzo et al., 2016; Sonter et al., 2016; van Zanten et al., 2016). These photos are used as an indicator for the revealed preferences of the general public. Despite the limitations of the approach such as a biased user population and behavior (Ruths and Pfeffer, 2014; Yoshimura and Hiura, 2017), previous studies using geotagged photos from the Flickr database have shown that the visitation rate extracted from the Flickr photos and user information matched well with the one calculated from the empirical visitor data (Wood et al., 2013; Keeler et al., 2015; Sonter et al., 2016). This highlights the reliability of the indicator to assess the demand for outdoor recreation and landscape aesthetics. While different photo repositories attract different user communities, van Zanten et al. (2016) found a high degree of correspondence among three photo repositories (i.e., *Flickr*, *Instagram*, and *Panoramio*). As spatially explicit information is a prerequisite for a better understanding of CES provision (Crossman et al., 2013; Brown and Fagerholm, 2015), geotagged photos provide an important opportunity to quantify and map CES (Weyand et al., 2016).

Previous studies using geotagged photos in CES analyses can be grouped into three categories. The first group focuses on the spatial and temporal information of photos (Casalegno et al., 2013; Keeler et al., 2015; Gliozzo et al., 2016; Tieskens et al., 2017). The focus of these studies has been on the location and the users by whom the photos were taken and uploaded. The Integrated Valuation of Ecosystem Services and Tradeoffs (InVEST) recreation model has applied the concept of photo-user-days (Sharp et al., 2016), which represents the total number of days in each mapping unit where at least one photo was taken by a user (Wood et al., 2013). The InVEST recreation model has begun to be applied to various CES analyses (Keeler et al., 2015; Sonter et al., 2016). A second group of the studies aims at relating landscape context and biophysical settings with the locations of geotagged photos (Pastur et al., 2016; Tenerelli et al., 2016; van Zanten et al., 2016; Oteros-Rozas et al., 2017). Pastur et al. (2016), for example, related the location of the photos representing the aesthetic value of the landscape of Southern Patagonia to biophysical characteristics such as the presence of water bodies and vegetation types. A third group analyzes the content of the photos. The focus of the analysis has been not only on the spatial and temporal information of the photos but also on the thematic information such as 'what' users have taken and uploaded (Minin et al., 2015). Traditionally, CES are manually classified (Richards and Friess, 2015; Thiagarajah et al., 2015; Pastur et al., 2016; Oteros-Rozas et al., 2017). Since the manual labeling of photos is a labor-intensive task (Minin et al., 2015), it is only applicable for a relatively small number of photos. Richards and Friess (2015) stated that one person could process approximately 140 photos per hour. Such a manual labeling approach is not feasible for 'big data' such as the immense data available in public photo repositories.

In this study, we suggest a new framework in CES mapping based on automated content analysis instead of manual labeling of photo content, 'Mapping cultural ecosystem services 2.0'. The suggested approach allows the interpretation of large volumes of photos based on their content within a feasible time frame. It focuses on contents of photos based on automated tags. A tag is a label or an annotation that provides simple and direct information of objectives (Schmitz, 2006), and often associated with images. Tagging allows users to manage and to share their online resources through keywords (Cattuto et al., 2007; Anderson et al., 2008; Tisselli, 2010). Analyses of tags are widely used in image or multimedia annotations such as *Flickr*, *Instagram*, and *Youtube* (Schmitz, 2006; Cattuto et al., 2007; Anderson et al., 2008). While Flickr provides users with tag suggestions, tagging is not mandatory and strictly guided in Flickr, thus often leading to photos with no user-provided tags (Sigurbjörnsson and van Zwol, 2008; Tisselli, 2010). Different languages used in tagging (e.g., English: *mountain*, German: *Berg*) is another source of data inconsistency. To overcome these problems with user-provided tags, we suggest using automated tags based on image recognition algorithms. Recently, Richards and Tunçer (2018) showed a potential to use automated keywords to analyze the contents of photos based on five tags provided by Google Cloud Vision. We propose here an alternative approach that builds on the rich image content information provided by the cloud computing platform, *Clarifai*[1] (Goodfellow et al., 2016; Rusk, 2016), and that uses a social network approach to identify thematic clusters of photos.

This study aims at developing a new method to analyze the content of large volumes of photos and to derive indicators of socio-cultural usage of landscapes. We applied the approach in a regional case study in Germany. The objectives of this study are i) to identify users' activities based on the contents of photos estimated by the machine-learned tags – 'what' are in the photos; ii) to identify CES hotspots in the study area – 'where' users visited particularly for CES related themes, and iii) to analyze whether those hotspots were related with other geographical features (i.e., protected areas).

2. Material and methods

2.1. Study area

The study was conducted in the Mulde basin in the federal state of Saxony in Germany (Fig. 1). The Czech part of the basin (6.2% of the basin) was not included in the analysis. The basin is covered by a mosaic of agricultural and forest patches. The largest part of the basin is used for agricultural purposes: 53% of the area in Germany is covered with cropland, and 7% of the area is pasture. Forest covers 26% of the basin. Urban areas (10.2%) were excluded from the analysis since we focused on outdoor recreations outside of urban areas (LfULG, 2017).

The Ore mountains ("Erzgebirge" in German) located in the southern part of the Mulde basin (Fig. 1) are one of the most important tourist areas in Saxony (Landestourismusverband Sachsen e.V., 2015). The number of tourists who stay overnight has increased since 2004, and reached more than three million overnight stays per year (Landestourismusverband Sachsen e.V., 2015). The main purpose of traveling to the Ore mountains is 'nature' (60%) followed by 'hiking' (58%) as named in a survey by the tourist office of the mountains (TV Erzegebirge, 2014). Sports tourism such as winter sports (42%) and mountain biking (42%) obtained particular attention in this region as well.

2.2. Data collection

The methodological framework for the data collection and the following analyses is presented in Supplementary Fig. SF1.

2.2.1. Flickr photos

The data collection was performed on the second of January 2017 and covered all the geotagged photos from the study area taken and uploaded between the 1st of January 2005 and the 31st of December 2016. The geotagged photos were identified and acquired through the Flickr Application Programming Interface (API)[2] based on the location information of the photos. As the Flickr API does not allow to use a polygon as a search boundary, we implemented a custom download function to select photos exactly within the target polygons. For each

[1] https://www.clarifai.com.

[2] https://www.flickr.com/services/api.