

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), Kreuzteckbahnstr. 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*

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 social-cultural uses of landscapes. It is expected to scale well with large numbers of photos and to be easily transferable to different regions.

*Corresponding author: Tel.: +49 8821 183 161; Fax: +49 8821 735 73
Email address: heera.lee@kit.edu (Heera Lee)

Keywords: social media photo services, cloud computing, tagging, keywords network analysis, mapping ecosystem services, big data analysis

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 explicate 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 an assessment of the use of spatial data for CES. Yet surveys are still often 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 (2017) 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 objective of the study is 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).

¹<https://www.clarifai.com>

2. Material and methods

2.1. Study area

The study was conducted in the Mulde basin in the federal state of Saxony in Germany (Figure 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 (Figure 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 Figure 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 polygon, we created the smallest bounding box covering the polygon. The photos found in the bounding box were once again filtered by the polygon. We only considered publicly available photos. In addition to the photos, meta information such as user id, pictured date, and coordinates was collected. The Flickr photo repository is unsteady in the sense that photos and meta-information can be altered any time.

2.2.2. Automatic tagging

We used an image annotation engine developed by Clarifai. The image annotation engine of Clarifai uses machine-learning to recognize visual patterns of images and videos and assigns tags automatically using several large deep convolutional neural networks (DCNNs) (Clarifai, 2016; Zeiler and Fergus, 2014). The DCNNs approach is one of the state-of-the-art algorithms in image recognition (Yang and Hospedales, 2015; Guo et al., 2016; Rawat and Wang, 2017). Another option would have been the automatic tag suggestion algorithm provided by Flickr.

²<https://www.flickr.com/services/api>

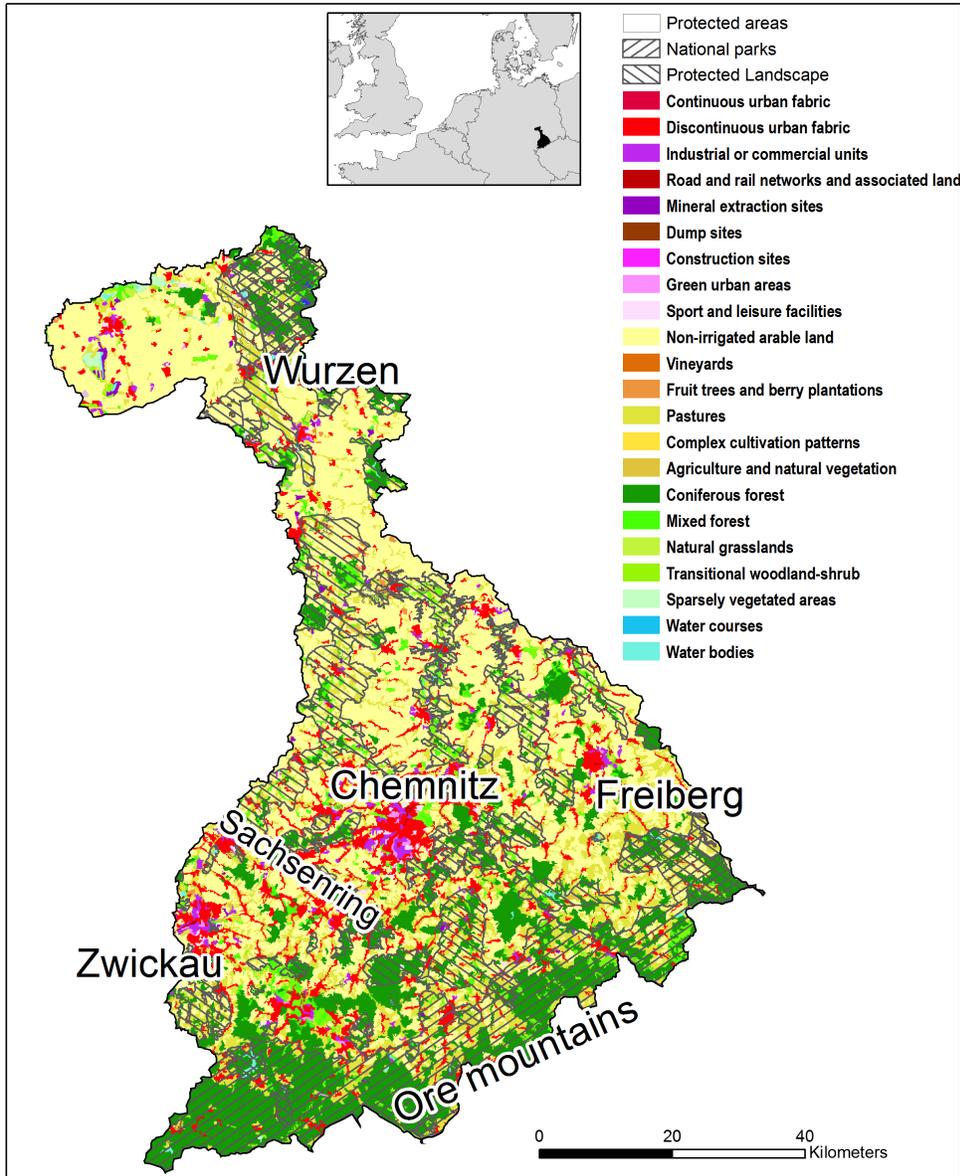


Figure 1: The Mulde basin in Saxony, Germany. Colors indicate different land cover types. Also, it shows the major urban areas in this region and the Ore mountains (“Erzgebirge” in German). The land cover data was taken from the 2006 CORINE Land Cover data (CLC2006; Umweltbundesamt, DLR-DFD 2009), and the data of special protection areas, special areas of conservation, nature reserves, protected landscapes and nature conservation parks provided by the Saxonian State Agency for Environment, Agriculture and Geology (LfULG) (LfULG, 2017).

At the time of the analysis no sufficient information on the image recognition algorithm used by Flickr was provided³ – therefore we decided to use the well-specified image annotation algorithm provided by Clarifai. For a detailed list of examples and the theory of deep learning

³source: <https://help.yahoo.com/kb/flickr/sln7455.html>

methods for visual recognition, we refer readers to the review studies by Guo et al. (2016) and Rawat and Wang (2017).

We used a pre-trained model (general-v1.3) provided by Clarifai to assign 20 tags per photo together with the associated probability of each tag; note that we used the term “tag” instead of the term “concept” in the Clarifai API document (Clarifai, 2016). We accessed Clarifai API version 1 through the GNU R (R Core Team, 2016) package `clarifai` (Sood, 2016).

2.3. Analysis

2.3.1. Cluster detection in tag-network

To identify themes of photos, we investigated networks of assigned tags based on tag co-occurrence. Networks analyses of semantic keywords have been applied in various research fields to group keywords with similar topics and to classify trending topics such as the analysis of scientific journal article keywords (Yi and Choi, 2012; Isenberg et al., 2017; Santonen and Conn, 2016), or the analysis of social networks such as Twitter (Lee et al., 2011).

We assumed that photos with similar themes share similar semantic tags. Tag co-occurrence is therefore regarded as an indicator to determine the contents of the photos. To analyze this, first we converted a 2-mode matrix (a photo-by-tags matrix) to a 1-mode tag co-occurrence matrix. The 1-mode matrix is a tag-by-tag matrix, in which the cells indicate how many times each pair of tags co-occurred in a photo (Schneegg and Bernard, 1996; Begelman, 2006; Anderson et al., 2008; Sigurbjörnsson and van Zwol, 2008; Hu et al., 2012; Mousselly-Sergieiev et al., 2013). Tags that appeared five times or more ($n = 1,316$) were included for further analysis because finding connections among those with fewer appearance would be illegible. Finally, we obtained a 1,316 by 1,316 co-occurrence matrix, which transformed into an undirected tag-network. We calculated eigenvector and degree centralities (Bonacich, 1987; Csardi and Nepusz, 2006) of each node (i.e., tag) to evaluate the importance of them.

We hierarchically clustered the tags and searched for the optimum number of the clusters. There exist a variety of algorithms for detecting clusters (i.e., often called communities) within networks (e.g., Fortunato et al., 2004; Luke, 2015; Yang et al., 2016). In this study, we used two widely-used algorithms and inter-compared the results: the Fast-greedy algorithm (Newman, 2004; Clauset et al., 2004) and the Walktrap algorithm (Pons and Latapy, 2005).

Fast-greedy algorithm The Fast-greedy algorithm is a clustering algorithm that merges clusters based on the expected improvement of modularity (see Section 2.3.1) for each pair of clusters (Newman, 2004). At each step the clusters are merged to maximally increase the modularity until no further improvement can be reached.

Walktrap algorithm The Walktrap algorithm is based on distance measured by random walks (Pons and Latapy, 2006). The key idea is that short random walks tend to stay in the closely connected parts (“trapped”), which can be considered as a cluster. This feature makes possible to efficiently identify clusters without having computationally intensive calculations based on eigenvectors (Pons and Latapy, 2005). The distance between node i and

r_{ij} is calculated based on the probability, P_{ij}^t , to go from node i to node j in a random walk of length t :

$$r_{ij} = \sqrt{\sum_{k=1}^n \frac{(P_{ik}^t - P_{jk}^t)^2}{d(k)}}, \quad (1)$$

where n is the number of nodes in the network and $d(j)$ is the degree of j . Note that the distance is adjusted by $d(j)$, as the random walk has a higher probability to go to nodes with many connections (i.e., high-degree). The distance between two clusters C_1 and C_2 of the network is defined similarly:

$$r_{C_1 C_2} = \sqrt{\sum_{k=1}^n \frac{(P_{C_1 k}^t - P_{C_2 k}^t)^2}{d(k)}} \quad (2)$$

with the probability P_{Cj}^t to go from cluster C to node j in t steps

$$P_{Cj}^t = \frac{1}{|C|} \sum_{i \in C} P_{ij}^t. \quad (3)$$

Ward’s hierarchical clustering method (Ward Jr, 1963) was used to merge clusters. Note that, in each step, only the adjacent clusters (i.e., having at least an edge between them) are merged to reduce computational time. Accordingly, it ensures that the merged clusters are connected to each other (Pons and Latapy, 2006).

Modularity The fitness of the detected clusters was evaluated by a criterion called ‘Modularity (Q)’ (Newman, 2004; Newman and Girvan, 2004), which has been widely used in many cluster detection algorithms (Fortunato et al., 2004; Danon et al., 2005; Pons and Latapy, 2006; Newman, 2006). The key idea is that the proportion of edges inside the same cluster (internal edges) is relatively high compared to the proportion of edges between the clusters if the clusters were appropriately assigned. Following Newman and Girvan (2004), we calculated Q as

$$Q = \sum_i (e_{ii} - a_i^2), \quad (4)$$

where k is the number of clusters and \mathbf{e} is a $k \times k$ symmetric matrix. An element of the matrix \mathbf{e} (e_{ii}) is the fraction of edges in the network that connects two clusters i and j . For a given division of a network into k clusters, the modularity measures the number of within cluster edges, relative to the expected value of random cluster edges with the same quantity in a cluster. If the number of within cluster edges is smaller than the random model, the modularity is close to zero. By increasing the number of clusters (k), one or two peaks of modularity are often found (Newman and Girvan, 2004). When the modularity is at its maximum peak, it is considered that the corresponding k is the optimal number of the clusters. For the robustness of the result, a Monte Carlo approach was used (Landau and Binder, 2005). For 100 iterations, we sampled 80% of the tags without replacement and

detected clusters, then calculated modularity to get the confidence interval of the modularity. The optimal cluster size was chosen based on the median value of the Monte Carlo modularity. The names of the clusters were assigned by the researchers based on the interpretation of the 10 tags with the highest eigenvector centrality in each cluster.

The tag network analysis was done in R version 3.3.1 (R Core Team, 2016) using the package `igraph` (Csardi and Nepusz, 2006) and `UserNetR` (Luke, 2017). The visualized position of the nodes in the network was determined by *Force Atlas2 multi-gravity algorithm* via the software Gephi (Bastian et al., 2009; Jacomy et al., 2014). We used the package `rgexf` (Yon et al., 2015) to transfer the R network data to Gephi.

2.3.2. Assigning dominant cluster to photos

Photos were assigned to a tag cluster based on a weighted majority voting. We assigned weights to tags based on the probability information supplied by the Clarifai engine for each tag to consider the credibility of the tags as well as to avoid ties. The proportion of the weighted tags belonging to the dominant cluster is the supporting ratio, which was calculated as

$$\text{Supporting ratio}(\%) = \max(W_{tags_k}) / \sum_k (W_{tags_k}) \times 100, \quad (5)$$

where W_{tags_k} is the weighted number of tags for each cluster and k is the index of clusters; The weighted W_{tags_k} equals to the number of the tags (W_{tags_k}) multiplied by the certainty of the tagging ($probs_k$) provided by the Clarifai engine. The higher the supporting ratio for a photo, the more tags belonging to the same cluster in the photo. The proportion of photos in each of the tag clusters were mapped at a 2.5 km grid pixel for visualization using R version 3.3.1 (R Core Team, 2016) packages `raster` (Hijmans, 2016) and `rgdal` (Bivand et al., 2017b).

Uncertainty assessment We considered two aspects of uncertainties: tagging uncertainty and dominant cluster assignment uncertainty. First, to evaluate the performance of automatic tagging, we manually analyzed tags for 635 randomly sampled photos (approx. 5% of the dataset). Compared to the human-interpreted tags, the machine-assigned tags were categorized into three groups: ‘relevant’, ‘possibly relevant’ (i.e., a tag is indirectly related to the content of the photo), and ‘irrelevant’. We assumed that ‘relevant’ and ‘possibly relevant’ are ‘correct’ and ‘irrelevant’ are ‘incorrect’. We calculated the Hamming loss (Tsoumakas et al., 2010) of the tagging which is an accuracy measure for multi-label classification. The Hamming loss ranges between zero (all tags are correct) and one (all tags are incorrect). Second, the uncertainty of the dominant cluster assignment was evaluated using the distributions of the supporting ratios for the dominant clusters.

2.3.3. Spatial distribution of photo-user-days

The photo-user-days were calculated as the total annual days that a photographer took at least one photo within a cell in the study area (Sharp et al., 2016). By using the photo-user-days instead of the total number of photos, one avoids a bias caused by an exceptionally high number of photos by a single photographer in a single day. Thus, this metric gives a good indication of the number of visitors to the area (Wood et al., 2013) in the study region during

the analyzed time period. In this study, we used the metric to give a spatial summary of the usage of the visited areas. Furthermore, we compared the photo-user-days calculated based on the whole data set (PUD_{all}) and the one calculated based on the CES-related photos (PUD_{ces}) to investigate a difference in the two maps. Herein CES related photos refer to the photos in clusters “landscape aesthetics” and “existence”. The patterns of PUD_{all} and PUD_{ces} were quantified by Pearson’s correlation.

To further evaluate the spatial information of the Flickr-photo-based indicator, we used aggregated information on the location of special protection areas, special areas of conservation, nature reserves, protected landscapes and nature conservation parks provided by the Saxonian State Agency for Environment, Agriculture and Geology (LfULG) (LfULG, 2017). Note that we combined the detailed protection-classes into a generic “protected area” class. The protected area polygons were intersected with the 1×1 km raster grid following the method suggested in Seo et al. (2016) which resulted in the fractional map of the protected area. We tested for the influence of the presence of protected areas (predictor) on the two groups (PUD_{all} and PUD_{ces}) (response) using a quasi-poisson generalized linear model. The model was adopted to account for the observed over-dispersion in the underlying count process (Zuur et al., 2009). Prior to the main analysis, spatial autocorrelation of the residuals was tested by Moran’s I , which was insignificant; therefore we did not explicitly consider the autocorrelative terms. We used the R packages `raster` (Hijmans, 2016), `rgeos` (Bivand and Rundel, 2018), `sp` (Pebesma and Bivand, 2005), and `spdep` (Bivand et al., 2017a) in this step of the analysis.

3. Results

During the studied period (2005–2016), 12,635 photos were uploaded by 725 users within the study area. The average number of the uploaded photos per user was 17.43. The distribution was right-skewed: the maximum number of photos uploaded by an individual user was 1,620, whereas 259 users uploaded only a single photo during the whole period. During the studied period, 27.5% of the users posted photos of the Mulde basin in multiple years. The number of the uploaded photos increased over the study period with a sharp increase in 2010. The largest number of photos was taken in May ($n = 1,893$, 15%) followed by June ($n = 1,762$, 13.9%) and August ($n = 1,527$, 12.1%) (Supplementary Figure SF2). The number of user-provided tags was also right skewed: in our database, 2,555 photos (20.2%) had no user-provided tag, and 590 photos had only one user-provided tag. The average number of user-provided tags was 10.03, and the maximum number of user-provided tags assigned in our database was 74 (cf. the maximum number of tags allowed in Flickr is 75). This uneven and skewed distribution of user-provided tags made use of user-provided tags unreliable for our analysis and led to the use of automated tagging from Clarifai.

3.1. Photo contents analysis – tagging and tag co-occurrence network analysis

The Clarifai engine assigned 20 tags per photo (examples provided in Supplementary Figure SF3). In total, 2,317 unique tags were assigned to the 12,635 photos. The most frequently assigned tag was ‘no person’ (assigned to 9,445 photos, 74.8%) followed by ‘outdoors’ (assigned to 8,550 photos, 67.7%), ‘nature’ (assigned to 6,362 photos, 50.3%), and ‘landscape’ (assigned to 6,181 photos, 48.9%) (Supplementary Table ST1). The 1,316 tags

that appeared more than five times were included in the tag-network analysis. For the 635 randomly sampled photos, the average Hamming loss of the assigned tags, i.e., the ratio of the mislabeled tags, was 0.196, which can be interpreted as 80.4% accuracy. The lowest Hamming loss was found for photos associated with the cluster “landscape aesthetics” (0.13), whereas the highest Hamming loss was associated with the photos related to the cluster “cultural landmark” (0.39). The tagging accuracy was higher in photos in CES-related clusters (“landscape aesthetics” and “existence”) compared to other clusters (Supplementary Table [ST2](#)).

The cluster composition with the Fast-greedy algorithm did not change with the increasing number of clusters larger than four (Supplementary Figure [SF4](#)). This is in line with results of [Yang et al. \(2016\)](#) that showed that the Fast-greedy algorithm underestimates the number of subgroups, whereas the Walktrap algorithm is more accurate regardless the network size - even for networks with more than 1000 nodes. Based on this result we continued using the Walktrap algorithm for the tag network analysis.

The co-occurrence of tags in photos was represented in a 1-mode co-occurrence matrix. The most frequent pair was ‘no person’ and ‘outdoors’ shown in 58.8% of the photos ($n = 7,428$), followed by the pairs of ‘nature’ and ‘outdoor’ (45.2%; $n = 5,713$), ‘nature’ and ‘no person’ (44.7%; $n = 5,642$), and ‘landscape’ and ‘no person’ (43.5%; $n = 5,499$) (Table [ST1](#)). The mean and the median degrees (i.e., the number of links) of the tag-network were 146.1 and 88, respectively. In Figure [2](#), we visualize the tag network with tags with higher eigenvector centrality located in the center. Modularity of the tag network increased sharply with increasing number of tag clusters up to nine ($Q = 0.193$), then gradually decreased (Supplementary Figure [SF5](#)). The number of clusters (k) was fixed as nine in the following analyses. The number of tags per cluster was unevenly distributed (Figure [2](#) and Table [1](#)): 296 tags in the largest cluster and 31 in the smallest cluster.

Clusters were denominated based on 10 tags with the highest eigenvector centrality (dominantly appeared tags) in each cluster (Table [1](#)). Among the nine clusters, two clusters were related to CES: “landscape aesthetics” and “existence”. The “landscape aesthetics” cluster included tags representing scenery, whereas the “existence” cluster included tags for specific species such as ‘butterflies’ or ‘flower’. The other clusters involved site-specific activities, but were related to non-CES activities such as “car racing” or “concerts”. An increase in the number of clusters can be used to reveal even more specific groups (Figure [3](#)).

Particularly high proportions of the CES clusters, “landscape aesthetics” and “existence”, were identified in proximity to the Ore mountains in the southern part of the basin (Figure [4](#)). This area is well known for outdoor activities ([Hodeck and Hovemann, 2016](#)). The proportion of the “car racing” cluster was high in the western part of the basin where the popular racing circuit ‘Sachsenring’ is located between Zwickau and Chemnitz (Figure [1](#)). Proportions of “equipment” and “transportation” clusters were high in the northern part of the basin.

Photos were frequently assigned tags from multiple clusters. Regarding the dominant cluster, 65.1% of the photos belonged to CES-related clusters: 53.7% of the photos to the cluster “landscape aesthetics” and 11.4% to “existence” (Figure [3](#)). The dominant clusters of the remaining 34.9% of the photos were not related with CES. The supporting ratio of the dominant cluster is the largest proportion in each photo (Eq. [5](#)). The highest supporting ratio was for the cluster “car racing” (60-80%) (Supplementary Figure [SF7](#)). Photos from the CES-related clusters had mostly a supporting ratio in the range of 40-60%; these photos

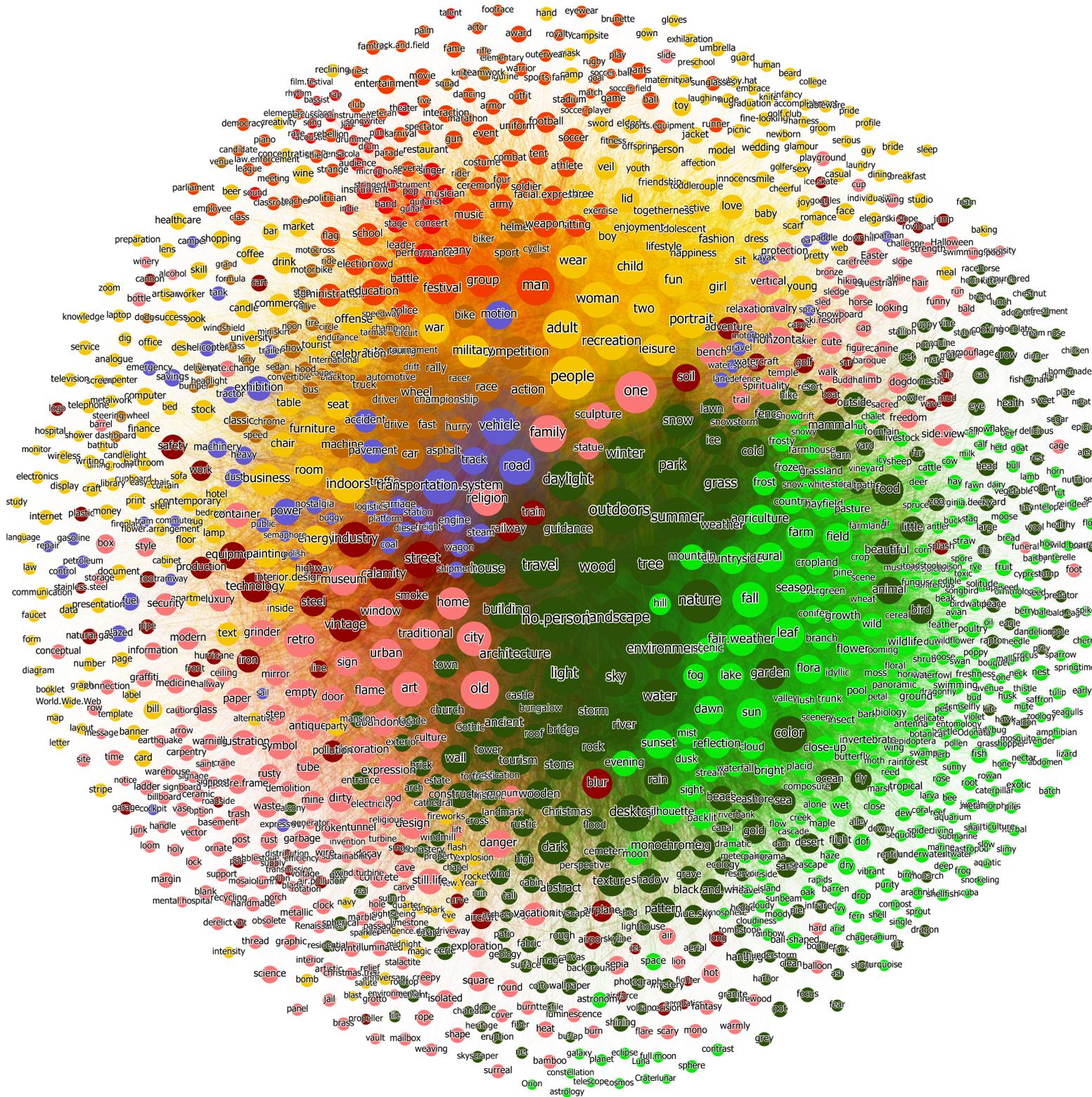


Figure 2: The tag-network of the Flickr photos. Nodes represent the machine-learned tags ($n = 1,316$) and links represent co-occurrence of tags in the photos. The size of a node and a label is scaled by eigenvector centrality, i.e., larger nodes were higher in centrality in the tag-network (Supplementary Table ST1). The colors indicate the tag clusters determined by Walktrap algorithm ($k = 9$): (a) dark green: landscape aesthetic, (b) green: existence, (c) dark yellow: indoors/people, (d) brown: car racing, (e) orange: festival, (f) pink: cultural landmark, (g) purple: transportation, (h) red: concerts, (i) dark red: equipment; this color scheme is applied to all following graphics.

Table 1: The top 10 tags with high eigenvector centrality in each cluster. The clusters “landscape aesthetics” and “existence” are considered as CES-related. Numbers in the parentheses are the total number of the tags of the clusters.

landscape aesthetic (248)	existence (296)	indoors/people (230)	car racing (66)	festival (92)
no person	leaf	people	fast	man
outdoors	flora	adult	drive	group
nature	fall	indoors	wheel	festival
wood	season	woman	race	education
summer	fair weather	portrait	driver	many
tree	bright	wear	car	election
travel	wild	child	hurry	battle
landscape	growth	room	action	music
grass	sun	girl	competition	administration
daylight	rural	furniture	machine	school
cultural landmark (231)	transportation (56)	concert (31)	equipment (66)	
old	transportation system	band	industry	
art	vehicle	performance	technology	
one	engine	musician	equipment	
family	road	pop	steel	
retro	track	concert	street	
religion	steam	singer	pollution	
traditional	power	instrument	vintage	
design	diesel	guitarist	production	
antique	shipment	guitar	train	
museum	carriage	stringed instrument	calamity	

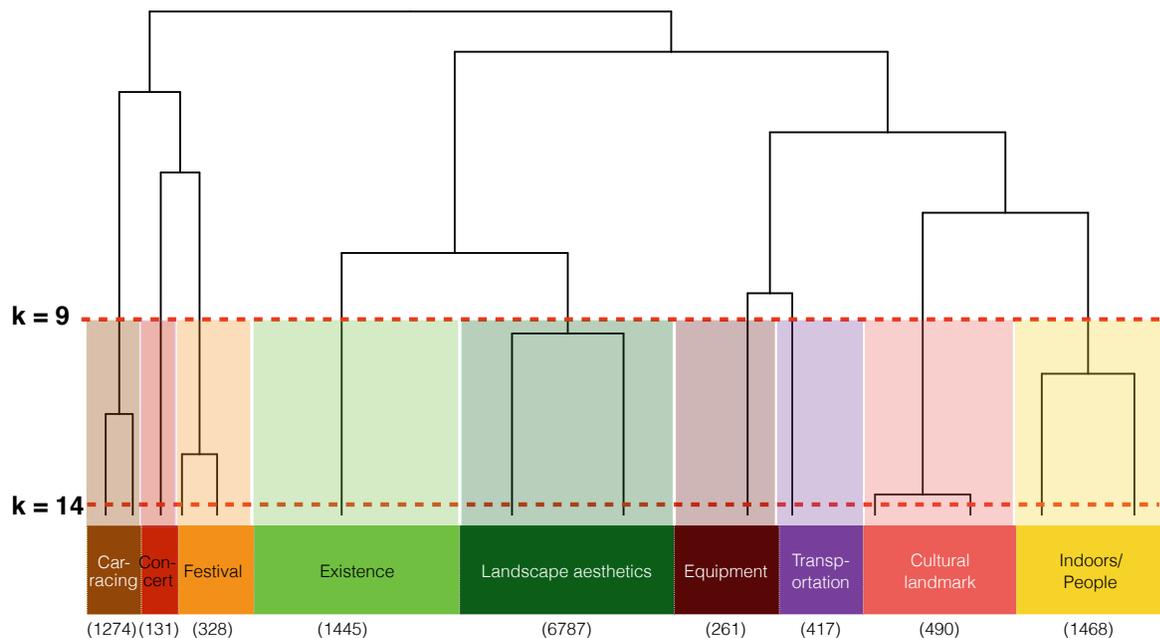


Figure 3: The dendrogram of the hierarchically clustered tags cut at the number of clusters (k) is 14. The dashed line in the middle and the colors in the bottom line and the boxes refer to the nine clusters ($k=9$) as used in this study. Numbers in the parentheses are the total number of the photos of the clusters.

may have been involved in multiple clusters (e.g., a photo with a car on the unpaved road might involve trees and therefore get also tags from the “landscape aesthetics” cluster).

The correlation between the photo-user-days of the total photos and the CES photos at the 1×1 km resolution was high ($r = 0.92$, Supplementary Figure SF6). However, the correlation differed strongly across the different CES types: the “landscape aesthetics” cluster was strongly correlated with the number of total photos ($r = 0.91$), while the “existence” cluster showed a different distribution ($r = 0.51$, Supplementary Figure SF6).

The presence of protected areas had a significant positive influence on the photo-user-days for both “existence” and “landscape aesthetics” (Supplementary Table ST3). Interestingly, the presence of protected areas had also a significant positive influence on the photo-user-days of the non-CES related photos. The number of photo-user-days increased with the share of protected areas around the location. The explanatory power of protected areas for photo-user-days measured by the explained deviance was however low for all clusters: “landscape aesthetics” (5.5%), “existence” (1.1%), and non-CES related photos (1.7%).

4. Discussion

Geotagged photos obtained from social media data have received increasing attention in identifying CES hotspots (Wood et al., 2013; Pastur et al., 2016; Tenerelli et al., 2016). However, the content of the photos has been so far often disregarded due to time and

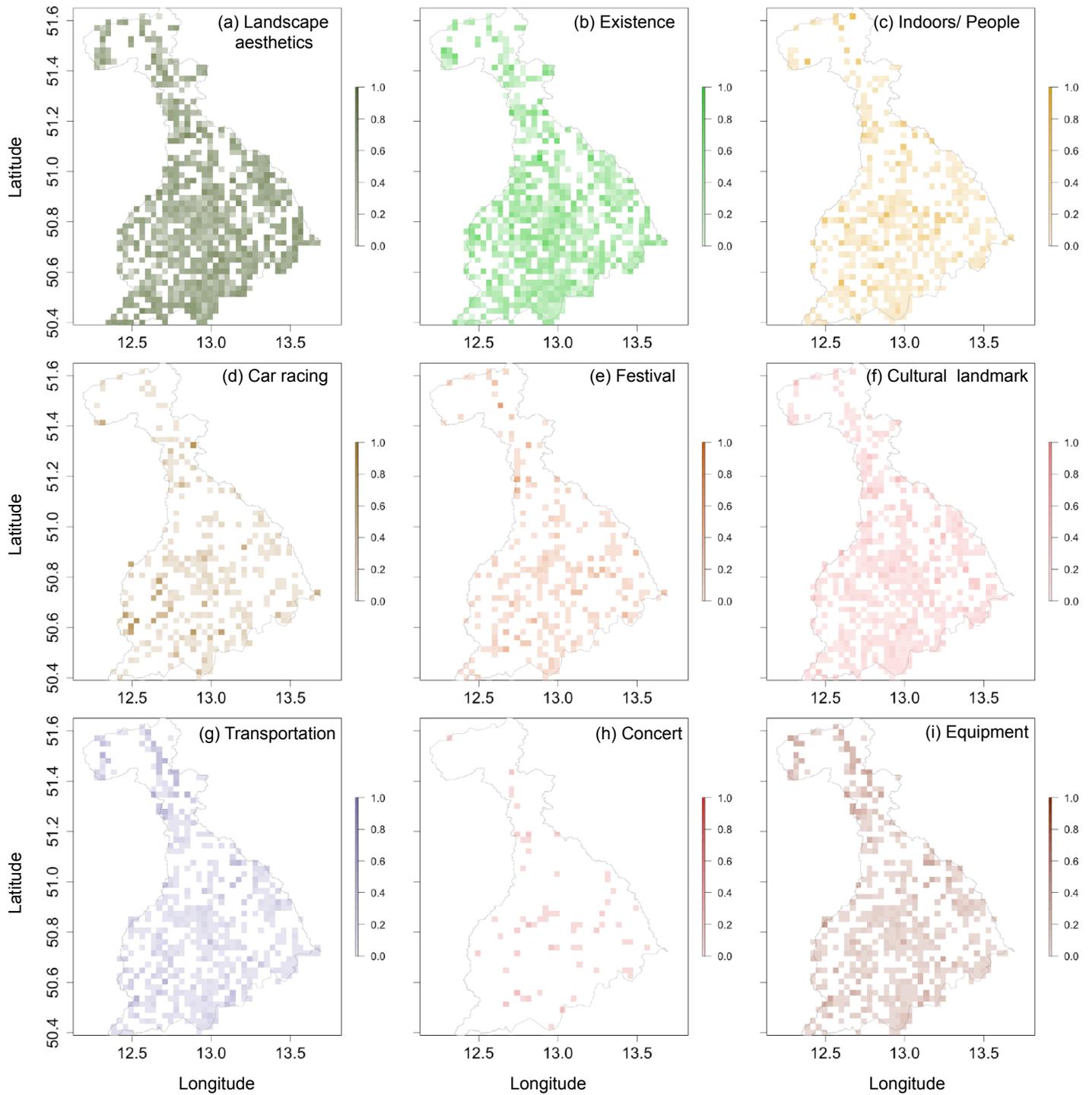


Figure 4: The average proportion of each cluster ($k = 9$) in the photos per cell ($2.5 \text{ km} \times 2.5 \text{ km}$). The more vivid colors indicate the higher proportion of a cluster in the photos of the cells.

resource limitations (Richards and Friess, 2015; Thiagarajah et al., 2015; Pastur et al., 2016; Richards and Tunçer, 2017). The combination of automatic tagging and tag-network analysis proposed in this study opens up a feasible way to analyze the contents of a large volume of photos. With the proposed method, we were able to distinguish detailed themes of photos (i.e., CES-related vs. non-CES-related themes (Table 1)) and to reproduce patterns in accordance with the ecological and cultural characteristics of the Mulde water basin in short time. This approach could be used to improve the commonly used ES evaluation tool, InVEST recreational model (Sharp et al., 2016), which calculates the photo-user-days based on the total number of geotagged photos with no consideration of the photo contents.

Content analysis of crowdsourced photos has a high potential to improve our understanding of the socio-cultural usage of the landscape. Despite this high potential, concerns remain regarding the representativeness of populations who uploaded photos in social media databases (Ruths and Pfeffer, 2014). However, given the limitations of geotagged real-time data collection within a limited time at a large scale (e.g., surveys and interviews) (Ford et al., 2016), the use of photo repositories offers an important opportunity to derive additional information on CES hotspots and detailed contents of those hotspots. Through content analysis, land managers can obtain not only information on hotspots where users frequently visit, but also on the activities performed by landscape users. This information allows land managers to respond accordingly, e.g. by restricting access to a specific vulnerable area or by providing facilities to steer the use of the landscape. Our approach offers a potential for managing landscapes used for outdoor recreation given a large number of crowdsourced photos available over the globe.

A major benefit of using automatic tagging is time efficiency in characterizing the contents of photos (Richards and Tunçer, 2017). Assigning 20 tags per photo for the 12,635 photos took approximately 3 hours, which is several orders of magnitude faster than the manual labeling (e.g., 140 photos per hour per researcher (Richards and Friess, 2015)). In our study, we used an externally trained classifier by machine-learning algorithms, instead of user-provided tags from Flickr since these were inconsistent and subjective. If we had used the Flickr tags, 20.2% of the photos would have been ignored because the users did not add tags to their photos. Furthermore, the unrestricted and non-standard characteristics of Flickr tagging (e.g., the number of tags per photo and language for tagging) potentially hinders interpretation of tags (Anderson et al., 2008; Tisselli, 2010). We, therefore, question the use of user-provided tags, e.g., for the pre-selection of photos (van Zanten et al., 2016).

However, there remained uncertainties in automatic tagging and cluster grouping. Manual validation of tags showed that automatic tagging is reliable - the 19.6% of tags were incorrect. Among photos, photos in the “landscape aesthetic” cluster showed the highest accuracy in automatic tagging. Furthermore, increasing tagging accuracy is expected due to the recent development of the research on artificial intelligence, for example, considering spatial and temporal contexts of photos (Weyand et al., 2016). Also, even though a single tag did not match with the photo, we assumed that the theme of the photos was represented by the combination of tags, not by a single tag. Co-occurrence of the tags ‘fall’ and ‘winter’, for instance, needs to be interpreted as “cold season of the year”.

For interpretation of tags, it should also be noticed that each tag can have multiple attributes assigned based on the results of the network analysis. For example, a tag, ‘tree’, is clustered in the “landscape aesthetic” cluster, but it does not automatically mean that

the tag, ‘tree’, cannot be used in other clusters. It means that the tag, ‘tree’, appeared more frequently with those such as ‘outdoors’ and ‘nature’ from the “landscape aesthetic” cluster rather than those such as ‘leaf’ and ‘flora’ from the “existence” cluster. The main difference between the “landscape aesthetics” and “existence” clusters in our dataset was the scale of photos. Photos included in the “existence” cluster were mainly macro photography (i.e., extreme close-up or zoom-in photography). We assumed that the close-up photography emphasizes the meaning of the specific subject more (so-called its “existence”) compared to the zoom-out photos. For photos in the CES clusters - that are all taken in natural environments - some overlap between the clusters seems logical. A photo of a butterfly taken outdoors would be likely to contain tags from both “landscape aesthetics” (e.g., nature) and “existence” clusters (e.g., butterfly, insect). Also, CES are interconnected (Daniel et al., 2012). The location of the split in the dendrogram is much below than the splits of the other lumped clusters in the hierarchical structure of clusters (Figure 3), which emphasizes that the two clusters shared similar features. Further research should consider how the scale of photos and the shared features should be dealt in CES research.

In our study, we identified nine clusters. The nine clusters selected in our case may not be directly applicable to other regions where different activities might be dominant. The decision about the number of clusters based on the co-occurrence of tags determines the thematic resolution of the further analysis. By increasing the number of clusters, one could get a more detailed list of CES as it breaks the large cluster into several sub-groups (Figure 3). If the number of the clusters in our case study was changed to fourteen instead of nine, the cluster “landscape aesthetics” would be further divided into two sub-groups (i.e., “agricultural land-based scenery” and “non-agricultural land-based scenery”), and the cluster “indoors/people” related with human activities, in general, was divided into “outdoor activities” and “indoors activities” (Figure 3). If the number of clusters was reduced to five, only a single CES would be present. Relying on a technical criterion such as modularity provides some guidance but needs to be thoroughly investigated.

Also, further research will analyze how strongly CES-related activities differ in space in relation to the landscape and socio-cultural settings. We compared the distribution of clusters with protected areas, but it can be further expanded with other landscape properties – such as terrain, a presence of water bodies, land use configuration but also accessibility and touristic infrastructure - that co-occur with these activities in space. It should also be explored how CES types differ with user groups such as tourists, elderly visitors, hikers, bikers, skiers, and visitors with kids. This information would enable a better representation of CES in landscape planning and spatially explicit trade-off analysis.

A limitation of the geotagged photos as an indicator for CES assessments can, however, be the limited coverage of different types of CES. The identified importance of landscape aesthetics as the main content of photos taken is in line with the recognition of landscape scenic beauty as the most popular motivation for photography (Richards and Friess, 2015), and the observation that tourists tend to value especially aesthetic values and recreation opportunities (Zoderer et al., 2016). We have distinguished different CES themes on the basis of unlabeled photos and shown that the spatial pattern of the different CES-related photos differs. However, we should note here that photo taking is limited in some recreation activities such as skiing (Wood et al., 2013; Tenerelli et al., 2016) and also other types of CES such as religious activities. Also, motivations for photos can differ even though a photo

contains similar features. A photo with a tree can represent the landscape aesthetic, but also the religious meaning depending on the society (Byers et al., 2001). Those hidden meaning could be difficult to be discovered by the content itself. Therefore, motivations for outdoor activities cannot be directly compared to the number of photos taken or the photo-user-days. Combinations with other methods to reveal CES can be beneficial when motivations for certain activities are questioned. There may be the potential need to link the geotagged photos based indicator to surveys. Thereby it should be possible to establish a better link and to quantify the reliability and the potential bias of our approach.

5. Conclusion

Geo-tagged photos in social media repositories such as Flickr provide a rich source of information on the perception and use of landscapes from a CES perspective. We were able to show that an analysis of the content of the photos allows to distinguish not only between non-CES and CES related photos but also to distinguish between different dimensions of CES. The combination of deep convolutional neural networks to tag photos with a social network analysis of the tag clouds led to reliable results. The presented approach is relatively fast, expected to scale well with large numbers of photos and to be easily transferable to different regions. It offers, therefore, a huge potential for the spatial explicit monitoring of CES activities at landscape and regional scale. An analysis of landscape properties in hotspot areas can provide additional information which landscape features affect the usability of the landscape. Further research should aim at quantifying the link between photo content and motivation to improve the reliability of the indicator.

Acknowledgements

This project was funded by the EU FP-7 project OPERAs (grant number 584 308393) and CONNECT from BiodiveERsA 2011 (BMBF: FK 01LC1103A). We acknowledge the support of the German Academic Exchange Service (DAAD) in the form of an International Travel Grant which enabled H.L. to attend the Natural Capital Symposium 2017. The work of B.S. was supported by Agenda Program of the Rural Development Administration (PJ00997802) in South Korea and Asian Institute for Energy, Environment & Sustainability, Seoul National University, South Korea. B.S. thanks to Dowon Lee who introduced network analysis for the first time.

Author contributions statement

H.L., B.S., T.K. and S.L. designed the research; H.L, B.S. and S.L. performed the research; H.L., B.S. and S.L. analyzed data; and H.L., B.S., T.K., and S.L. wrote the paper.

References

Anderson, A., Raghunathan, K., Vogel, A., 2008. TagEz: Flickr Tag Recommendation. Technical Report. Association for the Advancement of Artificial Intelligence.

- Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi : An open source software for exploring and manipulating networks, in: Third International AAAI Conference on Weblogs and Social Media, pp. 361–362.
- Begelman, G., 2006. Automated tag clustering: Improving search and exploration in the tag space, in: Proceedings of the Collaborative Web Tagging Workshop at WWW2005, Edinburgh, Scotland. pp. 15–33.
- van Berkel, D.B., Verburg, P.H., 2014. Spatial quantification and valuation of cultural ecosystem services in an agricultural landscape. *Ecological Indicators* 37, 163–174.
- Bivand, R., Altman, M., Anselin, L., Assunção, R., Berke, O., Bernat, A., Blanchet, G., Blankmeyer, E., Carvalho, M., Christensen, B., Chun, Y., Dormann, C., Dray, S., Halbersma, R., Krainski, E., Legendre, P., Lewin-Koh, N., Li, H., Ma, J., Millo, G., Mueller, W., Ono, H., Peres-Neto, P., Piras, G., Reder, M., Tiefelsdorf, M., Yu, D., 2017a. Spatial dependence: weighting schemes, statistics and models. Technical Report. version 0.6-9. <https://r-forge.r-project.org/projects/spdep/>.
- Bivand, R., Keitt, T., Rowlingson, B., 2017b. rgdal: Bindings for the Geospatial Data Abstraction Library. URL: <https://CRAN.R-project.org/package=rgdal>. R package version 1.2-8.
- Bivand, R., Rundel, C., 2018. rgeos: Interface to Geometry Engine - Open Source (GEOS). R package version 0.3-28.
- Bonacich, P., 1987. Power and centrality: A family of measures. *American journal of sociology* 92, 1170–1182.
- Brown, G., Fagerholm, N., 2015. Empirical PPGIS/PGIS mapping of ecosystem services: A review and evaluation. *Ecosystem Services* 13, 119–133.
- Byers, B.A., Cunliffe, R.N., Hudak, A.T., 2001. Linking the conservation of culture and nature: A case study of sacred forests in Zimbabwe. *Human Ecology* 29, 187–218.
- Casalegno, S., Inger, R., DeSilvey, C., Gaston, K.J., 2013. Spatial covariance between aesthetic value & other ecosystem services. *PLoS ONE* 8, e68437. doi:[10.1371/journal.pone.0068437](https://doi.org/10.1371/journal.pone.0068437).
- Cattuto, C., Loreto, V., Pietronero, L., 2007. Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences of the United States of America* 104, 1461–1464. doi:[10.1073/pnas.0610487104](https://doi.org/10.1073/pnas.0610487104).
- Clarifai, 2016. <https://www.clarifai.com/technology>. Accessed Dec 2017.
- Clauset, A., Newman, M.E.J., Moore, C., 2004. Finding community structure in very large networks. *Physical Review E* 70. doi:[10.1103/PhysRevE.70.066111](https://doi.org/10.1103/PhysRevE.70.066111).

- Cord, A.F., Roeßiger, F., Schwarz, N., 2015. Geocaching data as an indicator for recreational ecosystem services in urban areas: Exploring spatial gradients, preferences and motivations. *Landscape and Urban Planning* 144, 151–162. doi:[10.1016/j.landurbplan.2015.08.015](https://doi.org/10.1016/j.landurbplan.2015.08.015).
- 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 blue print for mapping and modelling ecosystem services. *Ecosystem Services* 4, 4–14. doi:[10.1016/j.ecoser.2013.02.001](https://doi.org/10.1016/j.ecoser.2013.02.001).
- Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. *InterJournal Complex Systems*, 1695. URL: <http://igraph.org>.
- Daniel, T.C., Muhar, A., Arnberger, A., Aznar, O., Boyd, J.W., Chan, K.M.A., Costanza, R., Elmqvist, T., Flint, C.G., Gobster, P.H., Grêt-Regamey, A., Lave, R., Muhar, S., Penker, M., Ribe, R.G., Schauppenlehner, T., Sikor, T., Soloviy, I., Spierenburg, M., Taczanowska, K., Tam, J., von der Dunk, A., 2012. Contributions of cultural services to the ecosystem services agenda. *Proceedings of the National Academy of Sciences of the United States of America* 109, 8812–8819. doi:[10.1073/pnas.1114773109](https://doi.org/10.1073/pnas.1114773109).
- Danon, L., Díaz-Guilera, A., Duch, J., Arenas, A., 2005. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment* 09, P09008. doi:[10.1088/1742-5468/2005/09/P09008](https://doi.org/10.1088/1742-5468/2005/09/P09008).
- 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.
- Ford, J.D., Tilleard, S.E., Berrang-Ford, L., Araos, M., Biesbroek, R., Lesnikowski, A.C., MacDonald, G.K., Hsu, A., Chen, C., Bizikova, L., 2016. Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences of the United States of America* 113, 10729–10732. doi:[10.1073/pnas.1614023113](https://doi.org/10.1073/pnas.1614023113).
- Fortunato, S., Latora, V., Marchiori, M., 2004. Method to find community structures based on information centrality. *Physical Review E* 70, 056104. doi:[10.1103/PhysRevE.70.056104](https://doi.org/10.1103/PhysRevE.70.056104).
- Gee, K., Burkhard, B., 2010. Cultural ecosystem services in the context of offshore wind farming: A case study from the west coast of Schleswig-Holstein. *Ecological Complexity* 7, 349–358. doi:[10.1016/j.ecocom.2010.02.008](https://doi.org/10.1016/j.ecocom.2010.02.008).
- Gliozzo, G., Pettorelli, N., Haklay, M.M., 2016. Using crowdsourced imagery to detect cultural ecosystem services: a case study in south wales, UK. *Ecology and Society* 21, 6. doi:[10.5751/ES-08436-210306](https://doi.org/10.5751/ES-08436-210306).
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.

- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual understanding: A review. *Neurocomputing* 187, 27–48. doi:[10.1016/j.neucom.2015.09.116](https://doi.org/10.1016/j.neucom.2015.09.116).
- von Heland, J., Folke, C., 2014. A social contract with the ancestors: Culture and ecosystem services in southern Madagascar. *Global Environmental Change* 24, 251–264. doi:[10.1016/j.gloenvcha.2013.11.003](https://doi.org/10.1016/j.gloenvcha.2013.11.003).
- Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling. Technical Report. R package version 2.5-8. <https://CRAN.R-project.org/package=raster>.
- Hodeck, A., Hovemann, G., 2016. Motivation of active sport tourists in a German highland destination: a cross-seasonal comparison. *Journal of Sport & Tourism* 20, 335–348. doi:[10.1080/14775085.2016.1235988](https://doi.org/10.1080/14775085.2016.1235988).
- Hu, J., Wang, B., Liu, Y., Li, D.Y., 2012. Personalized tag recommendation using social influence. *Journal of Computer Science and Technology* 27, 527–540. doi:[10.1007/s11390-012-1241-0](https://doi.org/10.1007/s11390-012-1241-0).
- Isenberg, P., Isenberg, T., Sedlmair, M., Chen, J., Moller, T., 2017. Visualization as seen through its research paper keywords. *IEEE Transactions on Visualization and Computer Graphics* 23, 771–780. doi:[10.1109/TVCG.2016.2598827](https://doi.org/10.1109/TVCG.2016.2598827).
- Jacomy, M., Venturini, T., Heymann, S., Bastian, M., 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE* 9, e98679. doi:[10.1371/journal.pone.0098679](https://doi.org/10.1371/journal.pone.0098679).
- 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, 76–81. doi:[10.1890/140124](https://doi.org/10.1890/140124).
- La Rosa, D., Spyra, M., Inostroza, L., 2016. Indicators of cultural ecosystem services for urban planning: A review. *Ecological Indicators* 61, 74–89. doi:[10.1016/j.ecolind.2015.04.028](https://doi.org/10.1016/j.ecolind.2015.04.028).
- Landau, D., Binder, K., 2005. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, New York, NY, USA.
- Landestourismusverband Sachsen e.V., 2015. http://www.ltv-sachsen.de/de/Wirtschaftsfaktor/Das-Jahr-2015_1708.html?sid=iktEMQGBzJS1KGcBwaMIqHzzpuy5dwuK. Accessed in Feb 2017.
- Lee, K., Palsetia, D., Narayanan, R., Patwary, M.M.A., Agrawal, A., Choudhary, A., 2011. Twitter trending topic classification, in: *ICDMW '11 Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops*, IEEE Computer Society, Washington, DC, USA. pp. 251–258. doi:[10.1109/ICDMW.2011.171](https://doi.org/10.1109/ICDMW.2011.171).
- LfULG, 2017. <http://www.umwelt.sachsen.de/umwelt/natur/24701.htm>. Accessed in Feb 2017.

- Luke, D., 2017. UserNetR: Data sets for A User's Guide to Network Analysis in R. URL: <https://github.com/DougLuke/UserNetR>. R package version 2.10.
- Luke, D.A., 2015. A User's Guide to Network Analysis in R. Springer International Publishing. doi:[10.1007/978-3-319-23883-8](https://doi.org/10.1007/978-3-319-23883-8).
- Milcu, A.I., Hanspach, J., Abson, D., Fischer, J., 2013. Cultural ecosystem services: A literature review and prospects for future research. *Ecology and Society* 18, 44. doi:[10.5751/ES-05790-180344](https://doi.org/10.5751/ES-05790-180344).
- Minin, E.D., Tenkanen, H., Toivonen, T., 2015. Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science* 3, 1–6. doi:[10.3389/fenvs.2015.00063](https://doi.org/10.3389/fenvs.2015.00063).
- Mousselly-Sergieh, H., Egyed-Zsigmond, E., Döller, M., Gianini, G., Kosch, H., Pinon, J.M., 2013. Tag similarity in folksonomies, in: *INFORSID 2013*, pp. 319–334.
- Newman, M.E.J., 2004. Fast algorithm for detecting community structure in networks. *Physical Review E* 69, 066133. doi:[10.1103/PhysRevE.69.066133](https://doi.org/10.1103/PhysRevE.69.066133).
- Newman, M.E.J., 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America* 103, 8577–8582. doi:[10.1073/pnas.0601602103](https://doi.org/10.1073/pnas.0601602103).
- Newman, M.E.J., Girvan, M., 2004. Finding and evaluating community structure in networks. *Physical Review E* 69, 026113. doi:[10.1103/PhysRevE.69.026113](https://doi.org/10.1103/PhysRevE.69.026113).
- Norton, L., Inwood, H., Crowe, A., Baker, A., 2012. Trialling a method to quantify the 'cultural services' of the English landscape using countryside survey data. *Land Use Policy* 29, 449–455. doi:[10.1016/j.landusepol.2011.09.002](https://doi.org/10.1016/j.landusepol.2011.09.002).
- Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., Plieninger, T., 2017. Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators* In Press. doi:[10.1016/j.ecolind.2017.02.009](https://doi.org/10.1016/j.ecolind.2017.02.009).
- Pastur, G.M., Peri, P.L., Lencinas, M.V., García-Llorente, M., Martín-López, B., 2016. Spatial patterns of cultural ecosystem services provision in Southern Patagonia. *Landscape Ecology* 31, 383–399. doi:[10.1007/s10980-015-0254-9](https://doi.org/10.1007/s10980-015-0254-9).
- Pebesma, E., Bivand, R., 2005. Classes and methods for spatial data in R. Technical Report 5 (2). R News. <http://cran.r-project.org/doc/Rnews/>.
- 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. doi:[10.1016/j.landusepol.2012.12.013](https://doi.org/10.1016/j.landusepol.2012.12.013).

- Pons, P., Latapy, M., 2005. Computing communities in large networks using random walks, in: pInar Yolum, G'ung'or, T., G'urgen, F., 'Ozturan, C. (Eds.), Computer and Information Sciences - ISCIS 2005. ISCIS 2005. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. volume 3733, pp. 284–293. doi:[10.1007/11569596_31](https://doi.org/10.1007/11569596_31).
- Pons, P., Latapy, M., 2006. Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications* 10, 191–218.
- R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: <http://www.R-project.org/>.
- Rawat, W., Wang, Z., 2017. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation* 29. doi:[10.1162/NECO_a_00990](https://doi.org/10.1162/NECO_a_00990).
- 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. doi:[10.1016/j.ecolind.2015.01.034](https://doi.org/10.1016/j.ecolind.2015.01.034).
- Richards, D.R., Tunçer, B., 2017. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services* in press. doi:[10.1016/j.ecoser.2017.09.004](https://doi.org/10.1016/j.ecoser.2017.09.004).
- Rusk, N., 2016. Deep learning. *Nature Methods* 13, 35. doi:[10.1038/nmeth.3707](https://doi.org/10.1038/nmeth.3707).
- Ruths, D., Pfeffer, J., 2014. Social media for large studies of behavior. *Science* 346, 1063–1064. doi:[10.1126/science.346.6213.1063](https://doi.org/10.1126/science.346.6213.1063).
- Santonen, T., Conn, S., 2016. Social network analysis based keyword analysis of ISPIM research topics, in: Proceedings of the XXVII ISPIM Innovation Conference, Blending Tomorrow's Innovation Vintage, Porto, Portugal.
- Schmitz, P., 2006. Inducing ontology from Flickr tags, in: Proceedings of the Workshop on Collaborative Tagging at WWW2006, Edinburgh, Scotland.
- Schnegg, M., Bernard, H.R., 1996. Words as actors: A method for doing semantic network analysis. *Field Methods* 8, 7–10. Sage Publications.
- Seo, B., Bogner, C., Koellner, T., Reineking, B., 2016. Mapping fractional land use and land cover in a monsoon region: The effects of data processing options. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* PP, 1–16. doi:[10.1109/JSTARS.2016.2544802](https://doi.org/10.1109/JSTARS.2016.2544802).
- Sharp, R., Tallis, H., Ricketts, T., Guerry, A., Wood, S., Chaplin-Kramer, R., Nelson, E., Ennaanay, D., Wolny, S., Olwero, N., Vigerstol, K., Pennington, D., Mendoza, G., Aukema, J., Foster, J., Forrest, J., Cameron, D., Arkema, K., Lonsdorf, E. and Kennedy, C., Verutes, G., Kim, C., Guannel, G., Papenfus, M., Toft, J., Marsik, M., Bernhardt, J., Griffin, R., Glowinski, K., Chaumont, N., Perelman, A., Lacayo, M., Mandle, L., Hamel, P., Vogl, A., Rogers, L., Bierbower, W., 2016. InVEST+ VERSION + User's Guide. Technical Report. The Natural Capital Project, Stanford University, University of Minnesota, The Nature Conservancy, and World Wildlife Fund.

- Sigurbjörnsson, B., van Zwol, R., 2008. Flickr tag recommendation based on collective knowledge, in: Proceedings of the 17th international conference on World Wide Web, New York, NY, USA. pp. 327–336. doi:[10.1145/1367497.1367542](https://doi.org/10.1145/1367497.1367542).
- Sonter, L.J., B. Watson, K., A. Wood, S., Ricketts, T.H., 2016. Spatial and temporal dynamics and value of nature-based recreation, estimated via social media. PLoS ONE 11, e0162372. doi:[10.1371/journal.pone.0162372](https://doi.org/10.1371/journal.pone.0162372).
- Sood, G., 2016. clarifai: R Client for the Clarifai API. URL: <http://github.com/soodoku/clarifai>. r package version 0.4.0.
- 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. doi:[10.1016/j.ecolind.2015.12.042](https://doi.org/10.1016/j.ecolind.2015.12.042).
- 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, 666–677. doi:[10.1007/s13280-015-0647-7](https://doi.org/10.1007/s13280-015-0647-7).
- Tieskens, K.F., Schulp, C.J., Levers, C., Kuemmerle, J.L.T., Plieninger, T., Verburg, P.H., 2017. Characterizing European cultural landscapes: Accounting for structure, management intensity and value of agricultural and forest landscapes. Land Use Policy 62, 29–39. doi:[10.1016/j.landusepol.2016.12.001](https://doi.org/10.1016/j.landusepol.2016.12.001).
- Tisselli, E., 2010. think Flickr think: A case study on strategic tagging. Communications of the ACM 53, 141–145. doi:[10.1145/1787234.1787270](https://doi.org/10.1145/1787234.1787270).
- Tsoumakas, G., Katakis, I., Vlahavas, I., 2010. Mining Multi-label Data. Springer US, Boston, MA. URL: http://dx.doi.org/10.1007/978-0-387-09823-4_34, doi:[10.1007/978-0-387-09823-4_34](https://doi.org/10.1007/978-0-387-09823-4_34).
- TV Erzgebirge, 2014. Destinationsstrategie Erzgebirge [Destination strategy of the Ore Mountains]. Technical Report. Annaberg-Buchholz: Tourismusverband Erzgebirge e.V.
- Ward Jr, J.H., 1963. Hierarchical grouping to optimize an objective function. Journal of the American statistical association 58, 236–244.
- Weyand, T., Kostrikov, I., Philbin, J., 2016. Planet-photo geolocation with convolutional neural networks, in: European Conference on Computer Vision, Springer. pp. 37–55.
- 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. doi:[10.1038/srep02976](https://doi.org/10.1038/srep02976).
- Yang, Y., Hospedales, T.M., 2015. Deep neural networks for sketch recognition. arXiv:1501.07873v1 .
- Yang, Z., Algesheimer, R., Tessone, C.J., 2016. A comparative analysis of community detection algorithms on artificial networks. Scientific Reports 6, 30750. doi:[10.1038/srep30750](https://doi.org/10.1038/srep30750).

- Yi, S., Choi, J., 2012. The organization of scientific knowledge: the structural characteristics of keyword networks. *Scientometrics* 90, 1015–1026. doi:[10.1007/s11192-011-0560-1](https://doi.org/10.1007/s11192-011-0560-1).
- Yon, G.V., Lacoa, J.F., Kunst, J.B., 2015. rgexf: Build, Import and Export GEXF Graph Files. URL: <https://CRAN.R-project.org/package=rgexf>. R package version 0.15.3.
- 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.
- van Zanten, B.T., van Berkelan, 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, 12974–12979. doi:[10.1073/pnas.1614158113](https://doi.org/10.1073/pnas.1614158113).
- Zeiler, M.D., Fergus, R., 2014. Visualizing and understanding convolutional networks, in: *Computer Vision - ECCV 2014*, Springer International Publishing, Switzerland. pp. 818–833. doi:[10.1007/978-3-319-10590-1_53](https://doi.org/10.1007/978-3-319-10590-1_53).
- Zoderer, B.M., Tasser, E., Erb, K.H., Stanghellini, P.S.L., Tappeiner, U., 2016. Identifying and mapping the tourists' perception of cultural ecosystem services: A case study from an Alpine region. *Land Use Policy* 56, 251–261. doi:[10.1016/j.landusepol.2016.05.004](https://doi.org/10.1016/j.landusepol.2016.05.004).
- Zuur, A., Ieno, E., Walker, N., Saveliev, A., Smith, G., 2009. *Mixed Effects Models and Extensions in Ecology with R*. Springer.