



Integrating remote sensing and social media data advances assessment of cultural ecosystem services

Oleksandr Karasov¹, Stien Heremans^{2,3}, Mart Külvik⁴, Artem Domnich⁵, Iuliia Burdun⁶, Ain Kull⁷, Aveliina Helm⁷, and Evelyn Uuemaa⁷

¹Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland (oleksandr.karasov@helsinki.fi)

²Research Institute for Nature and Forest (INBO), Brussels, Belgium

³Department of Earth and Environmental Sciences, KU Leuven, Leuven, Belgium

⁴Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Tartu, Estonia

⁵Institute of Computer Science, University of Tartu, Tartu, Estonia

⁶Department of Built Environment, Aalto University, Espoo, Finland

⁷Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia

Over the past decade, we witnessed a rapid growth in the use of social media data when assessing cultural ecosystem services (CESs), like modelling the supply-demand relationships. Researchers increasingly use user-generated content (predominantly geotagged pictures and texts from Flickr, Twitter, VK.com) as a spatially explicit proxy of CES demand. However, for modelling CES supply most of such studies relied on simplistic geospatial data, such as land cover and digital elevation models. As a result, our understanding of the favourable environmental conditions underlying good landscape experience remains weak and overly generic.

Our study aims to detect the spatial disparities between population density and CES supply in Estonia in order to prioritise them for further in-depth CES assessment and green and blue infrastructure improvements. We relied on Flickr and VK.com photographs to detect the usage of three CESs: passive landscape watching, active outdoor recreation, and wildlife watching (biota observations at organism and community levels) with automated image content recognition via Clarifai API and subsequent topic modelling. Then, we used Landsat-8 cloudless mosaic, digital elevation and digital surface models, as well as land cover model to derive 526 environmental variables (textural, spectral indices and other indicators of landscape physiognomy) via the Google Earth Engine platform. We conducted an ensemble environmental niche modelling to analyse the relative strength and directions of relationships between these predictors and the observed occurrence of CES demand. Based on multicollinearity and relative importance analysis, we selected 21 relevant and non-collinear indicators of CES supply. With these indicators as inputs, we then trained five models, popular in environmental niche modelling: Boosted Regression Trees, Generalized Linear Model, Multivariate Adaptive Regression Spline, Maxent, and Random Forest. Random Forest performed better than the other models for all three CES types, with the average 10-fold cross-validation area under curve > 0.9 for landscape watching, >0.87 for outdoor recreation, and >0.85 for wildlife watching. Our modelling allowed us to estimate the share of the

Estonian population residing in the spatial clusters of systematically high and low environmental suitability for three considered CESs. The share of the population residing in the clusters of low environmental suitability for landscape watching, outdoor recreation, and wildlife watching is 5.5%, 3.1%, and 7.3%, respectively. These results indicate that dozens of thousands of people in Estonia (population is >1.3 million) likely have fewer opportunities for everyday usage of considered CESs. However, these results are biased as there was not enough evidence in social media for CES use in some of these areas.

Although our results should be treated with caution, because social media data are likely to contain a considerable sampling bias, we have demonstrated the added value of remote sensing data for CES supply estimation. Given nearly global and continuously updated satellite imagery archives, remote sensing opens new perspectives for monitoring the loss and gains in landscape suitability for CES across temporal and spatial scales. As such, we can better account for the intangible underlying geospatial features that can influence economic and environmental decision-making.