Remote Sensing of Environment 216 (2018) 482–496

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

# Generalizing machine learning regression models using multi-site spectral libraries for mapping vegetation-impervious-soil fractions across multiple cities

Akpona Okujeni[a,*], Frank Canters[b], Sam D. Cooper[a], Jeroen Degerickx[c], Uta Heiden[d], Patrick Hostert[a,e], Frederik Priem[b], Dar A. Roberts[f], Ben Somers[c], Sebastian van der Linden[a,e]

[a] *Geography Department, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*
[b] *Department of Geography, Cartography and GIS Research Group, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium*
[c] *Division of Forest, Nature and Landscape, KU Leuven, Celestijnenlaan 200E, 3001 Leuven, Belgium*
[d] *German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, 82234 Wessling, Germany*
[e] *Integrative Research Institute on Transformations of Human-Environment Systems, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*
[f] *Department of Geography, University of California, Santa Barbara, CA 93106, United States*

ARTICLE INFO

*Keywords:*
Urban remote sensing
Imaging spectrometry
EnMAP
Support vector regression
Land cover fraction mapping
Generalization

ABSTRACT

Forthcoming spaceborne imaging spectrometers will provide novel opportunities for mapping urban composition globally. To move from case studies for single cities towards comparative and more operational analyses, generalized models that may be transferred throughout space are desired. In this study, we investigated how single regression models can be spatially generalized for vegetation-impervious-soil (VIS) mapping across multiple cities. The combination of support vector regression (SVR) with synthetically mixed training data generated from spectral libraries was used for fraction mapping. We developed three local models based on separate spectral libraries from Berlin (Germany), Brussels (Belgium), and Santa Barbara (U.S.), and a generalized model based on a combined multi-site spectral library. To examine the performance and transferability of the generalized model compared to local models, we first applied all model variants to simulated Environmental Mapping and Analysis Program (EnMAP) data from the three cities that were represented in the models, i.e., known sites. Next, we transferred the models to two unknown sites not represented in the models, San Francisco Bay Area (U.S.) and Munich (Germany). In the first mapping constellation, results demonstrated that the generalized model was capable of accurately mapping VIS fractions across all three known sites. Average mean absolute errors (AV-MAEs) were 8.5, 12.2, and 11.0% for Berlin, Brussels, and Santa Barbara. The performance of the generalized model was very similar to the local models, with ΔAV-MAEs falling within a range of ± 0.7%. A detailed assessment of fraction maps and class-wise accuracies confirmed that modeling errors related to remaining limitations of urban mapping based on optical remote sensing data rather than to the choice between a local or generalized model. For the second mapping constellation, the generalized model proved to be useful for mapping vegetation and impervious fractions in the unknown sites. MAEs for both cover types were 5.4 and 10.9% for the San Francisco Bay Area, and 6.3 and 15.4% for Munich. In contrast, the three local models were only found to have similar accuracies as the generalized model for one of the two sites or for individual VIS categories. Despite the enhanced transferability of the generalized model to the unknown sites, deficiencies remained for accurate soil mapping. MAEs were 22.4 and 12.3%, and high over - and underestimations were observed at the low and high end of the fraction range. These shortcomings indicated possible limitations of the spectral libraries to account for the spectral characteristics of soils in the unknown sites. Overall, we conclude that the combination of SVR and synthetically mixed training data generated from multi-site libraries constitutes a flexible modeling approach for generalized urban mapping across multiple cities.

* Corresponding author.
*E-mail address:* akpona.okujeni@geo.hu-berlin.de (A. Okujeni).

https://doi.org/10.1016/j.rse.2018.07.011
Received 28 November 2017; Received in revised form 6 July 2018; Accepted 8 July 2018

## 1. Introduction

Optical satellite remote sensing has great potential for characterizing urban environments. Particularly the mapping of urban composition according to the vegetation-impervious-soil (VIS) framework (Ridd, 1995) has received considerable attention. This framework represents urban areas in a continuous ternary mixing diagram of vegetation, impervious, and soil cover fractions, and was proposed as a standardized mapping scheme to support urban ecosystem analysis, to monitor urban change processes, and for linking urban composition to urban morphology (Ridd, 1995). Numerous studies have exploited multispectral satellite data for mapping individual or all VIS components for different cities around the world (Phinn et al., 2002; Powell et al., 2007; Pu et al., 2008; Rashed et al., 2003). To match the quantitative nature of the VIS framework and to account for mixed pixels typical of urban environments, sub-pixel cover fraction mapping techniques based on linear spectral mixture analysis and its variants (Deng and Wu, 2013; Rashed et al., 2003; Small and Lu, 2006), or based on quantitative empirical modeling with regression algorithms or neural networks (Okujeni et al., 2015; Pu et al., 2008; Walton, 2008) have been widely adopted.

With the ongoing development in spaceborne imaging spectrometry, novel opportunities for a variety of environmental research fields arise. Several hyperspectral satellite missions are currently in preparation, including the Environmental Mapping and Analysis Program (EnMAP, Guanter et al., 2015). EnMAP will frequently collect high spectral resolution images over large geographic areas at a 30 m resolution, which will be passed through standardized processing chains to obtain consistent surface reflectance products. The availability of such imagery will open up new opportunities for global comparative analyses of urban composition and compositional changes. The hyperspectral information will most likely enhance the spectral separability of urban construction materials and vegetation types (Gamba and Dell'Acqua, 2006; Herold et al., 2003), which can lead to more accurate VIS maps. In addition, the spectral benefits will most likely enable the extension of VIS mapping into thematically more detailed impervious and vegetation sub-categories (Okujeni et al., 2015; Roberts et al., 2012; Wetherley et al., 2017). By offering the opportunity to derive more accurate and detailed descriptions of urban composition worldwide, spaceborne imaging spectrometry has considerable potential to introduce a new quality to urban remote sensing.

With regard to satellite imagery with large spatial and high temporal coverage, more generalized mapping and monitoring approaches that can be transferred throughout space and/or time are desired. Generalizing methods in remote sensing were originally introduced as "signature extension" (Botkin et al., 1984). Signature extension aims to move beyond image-by-image approaches by applying a model, which was trained with signatures from one image, to another image from a different geographic location (spatial generalization), a different acquisition time (temporal generalization), a different sensor (across-sensor generalization), or a combination thereof (Foody et al., 2003; Pax-Lenney et al., 2001; Woodcock et al., 2001). In the urban context, globally applicable models were mainly developed for mapping urban extent from MODIS data (Schneider et al., 2010) or different built-up cover types using Landsat data in combination with a digital surface model (Pesaresi et al., 2016). Several studies have explored the use of generalized approaches for mapping urban composition from Landsat data only. Amongst others, Small (2005) demonstrated the use of a global spectral mixture model to map high albedo substrate, vegetation, and dark surface fractions for 28 cities worldwide. Kaspersen et al. (2015) mapped impervious fractions for eight European cities using a regionally generalized regression model. Sexton et al. (2013) made use of a temporally generalized regression model to produce annual maps of impervious surface cover in the Washington-Baltimore metropolitan region between 1984 and 2010. Wang et al. (2014) examined the temporal generalization capability of neural networks, random forest classification, and regression trees for characterizing vegetation fraction dynamics in Zhongwei City between 1990 and 2010.

The availability of stable and comparable surface reflectance units between images is a crucial factor for successfully extending training signatures across scenes (Woodcock et al., 2001). Differences in acquisition conditions and dates, sensor characteristics, reflectance retrieval algorithms, etc., can lead to substantial radiometric inconsistencies between images, which can limit the model transfer across scenes (Olthof et al., 2005; Pax-Lenney et al., 2001). Studies therefore have focused on adjusting images by improved radiometric and atmospheric pre-processing (Olthof et al., 2005; Pax-Lenney et al., 2001), or on adapting the training signatures to individual images (Gray and Song, 2013).

The availability of comprehensive training information is another important factor influencing model generalization. Training signatures may not represent the full diversity and variability of cover types present in new regions or at different periods of time, which limits the model transferability across space and time. The use of multi-source training information representative for multiple sites or dates constitutes a suitable means to overcome these limitations. In the urban context, for example, the aforementioned studies by Kaspersen et al. (2015) and Sexton et al. (2013) demonstrated the use of specific multi-site and multi-annual training areas with impervious reference fractions calculated from high resolution land cover information for spatial and temporal regression model generalization. While this approach is straightforward and produces reliable results, the regression model training fully depends on the availability of accurately co-registered, high resolution reference data. In this regard, multi-source spectral libraries constitute an alternative solution that is independent from the availability of representative training areas and therefore marks a step forward towards more generalized mapping. For example, Michishita et al. (2012) demonstrated the use of a combined multi-annual library for mapping changes in urban composition for five time steps between 1987 and 2009. Similarly, Dudley et al. (2015) demonstrated that a multi-seasonal library can improve vegetation mapping and may support assessments regardless of the seasonality of the input image. All of these studies share the conclusion that multi-source training information provides a suitable means to enhance the spatial and temporal applicability and transferability of mapping models, and consequently their generalization capabilities.

With regard to the unprecedented opportunities of forthcoming spaceborne imaging spectrometers for assessing urban composition globally, the overarching goal of this study was to investigate how single regression models can be spatially generalized for VIS fraction mapping across multiple cities. We selected the combination of support vector regression (SVR) with synthetically mixed training data for fraction mapping (Okujeni et al., 2013). This approach proved to effectively handle urban cover types with high within-class variability, between-class similarity, and spectral mixing from airborne to spaceborne scales (Okujeni et al., 2015; Rosentreter et al., 2017). Moreover, this library-based approach is independent from the spatial context of the work with training areas and appears better suited for more generalized regression model training. Based on spectral libraries for the cities of Berlin (Germany), Brussels (Belgium), and Santa Barbara (U.S.), we created three separate models from the single libraries (henceforth referred to as "local models"), and one model from a combined multi-site library (henceforth referred to as "generalized model"). VIS mapping was subsequently conducted on simulated