Humboldt-Universität zu Berlin – Geographisches Institut

MAPPING URBAN SURFACE MATERIALS USING IMAGING SPECTROSCOPY DATA

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

im Fach Geographie

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Chaonan Ji (M.Sc.)

Präsident (komm.) der Humboldt-Universität zu Berlin

Prof. Dr. Peter Frensch

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät

Prof. Dr. Elmar Kulke

Gutachter/in:
1. Prof. Dr. Tobia Lakes
2. Prof. Dr. Hannes Feilhauer
3. Prof. Dr. Angela Lausch
4. Prof. Dr. Sebastian van der Linden

Datum der Einreichung: 20.12.2021

Datum der Promotion: 07.04.2022

Declaration

I declare that I have completed the thesis independently using only the aids and tools specified. I have not applied for a doctor's degree in the doctoral subject elsewhere and do not hold a corresponding doctor's degree. I have taken due note of the Faculty of Mathematics and Natural Sciences PhD Regulations, published in the Official Gazette of Humboldt-Universität zu Berlin no. 42/2018 on 11/07/2018.

Date: 04.01.2022

Signature:

Acknowledgements

The harvest of my doctoral time is not only this dissertation, but also the people I have met and the experiences I have gained on this journey. I would like to express my deep gratitude to all those who have contributed to this work and have accompanied and supported me along the way.

First and foremost, I would like to thank Prof. Tobia Lakes, my first supervisor, for her valuable advice and guidance along my way. And I would like to express my deepest gratitude to Prof. Hannes Feilhauer, as my second supervisor, whose reliable motivation, ideas, and support on the topics are invaluable. He always gives me broader understanding of topics, and therefore helped me to make this work possible. In addition, I am grateful for the thoughtful supervise and inspiring discussions of Dr. Uta Heiden. As my DLR internal supervisor, she has guided me to build the foundation of my hyperspectral knowledge, and continuously support each of my topics, including this thesis. Moreover, special thanks go to Prof. Angela Lausch and Prof. Sebastian van der Linden, who supported the work as referees.

In this particular pandemic situation, I am very fortunate to have several great mentors to keep me on track with the research. I would like to thank Dr. Marianne Jilge for her guidance, patience, and honest suggestions, which gave me a lot of confidence and rigorous research thinking at the beginning of my doctoral time. My heartfelt thanks go to Dr. Wieke Heldens for the great motivation, especially for taking care of me during the most stressful study period. I wish her massive success in her future work. Many thanks to Dr. Thomas Esch for his generous support and inspiring supervision. With his support, I was able to concentrate on the topics and had no other concerns. And my sincere thanks go to Dr. Martin Bachmann for his constant patience, valuable discussions, and comprehensive suggestions not only with the topics themselves but also with the presentation of the work. These mentors have been so important to each step of my journey that I could stumble without any of them along the way.

I would like to thank all my colleagues at DLR for their support and pleasant environment. Cordially thank Simone Zepp and Daniela Palacios Lopez for their constant support and encouragement during all stages of the doctoral period. Thank you to colleagues, Sina Starmans, Dr. Annekatrin Metz-Marconcini, Dr. Bennjamin Leutner, Martin Habermeyer,

Julian Zeidler, Dr. Mattia Marconcini, Elisabeth Brozoska, Dr. Nicole Pinnel, Stefanie Holzwarth, and Dr. David Marshall. It is a pleasure to work with your passion and warm support. In particular, I would like to thank Prof. Claudia Künzer for her support and for creating a pleasant study environment. I would also like to thank all the proofreaders, who invested their time and energy in this work.

I would also like to thank my parents for their loving support. 感谢你们的包容和支持，让我能够实现心之所向，素履以往。志之所趋，无远弗届。

Abstract

Urban environment and its processes directly affect human life. Detailed and up-to-date urban surface material maps are of great importance to modelers studying meteorology, climatology and ecology, as well as to authorities seeking to understand the urban growth dynamics and spatial evolution. However, mapping urban surface materials is challenging due to the complex spatial patterns. An established source of up-to-date information is remote sensing, as demonstrated by the widespread usage of SAR, LiDAR and optical data. Data from imaging spectrometers can identify detailed spectral features of surface materials through the fine and continuous sampling of the electromagnetic spectrum, which cannot be achieved with the same accuracy using multispectral or RGB images. To date, numerous studies in urban surface material mapping have been using data from airborne imaging spectrometers with high spatial resolution, demonstrating the potential and providing good results. Compared to these sensors, spaceborne imaging spectrometers have regional or global coverage, high repeatability, and avoid expensive, time-consuming, and labor-intensive flight campaigns. However, the spatial resolution of current spaceborne imaging spectroscopy data (also known as hyperspectral data) is about 30 m, resulting in a mixed pixel problem that is challenging to handle with conventional mapping approaches.

The main objective of this study is to perform urban surface material mapping with imaging spectroscopy data at different spatial scales, simultaneously explore the information content of these data to detect the chemical and physical properties of surface materials, and take the mixed-pixel problem into account. Specifically, this thesis aims to (1) map solar photovoltaic modules using airborne imaging spectroscopy data based on their spectral features; (2) investigate the sampling robustness of urban material gradients; (3) analyze the area transferability of urban material gradients.

To this end, we detected solar photovoltaics with an overall accuracy of about 80% to 90% by creating and combining spectral indices. This dissertation proved that the developed approach is suitable for accurate photovoltaic detection. We also demonstrated that the concept of urban surface material gradients is robust in sampling and transferable between similar urban areas. With these results, urban material gradients can be a generic technique for urban mapping with spaceborne imaging spectroscopy data. The methods developed in

the three parts of this dissertation improve the usefulness of imaging spectroscopy data for urban material detection from a classical method to the new concept of urban gradients, from airborne to spaceborne data, from pure pixel detection to solving the mixed pixel problem. By introducing and enhancing the gradient concept in urban mapping, the mixed pixel problem can be tackled, which is a promising approach for the analysis of imaging spectroscopy data from ongoing and upcoming spaceborne sensors. Overall, this thesis provides promising urban surface material mapping results by proposing a physical feature based approach as well as confirming and laying the foundation of the generic gradient concept in urban material studies. Further work can build on these results and could open a new field for the application of spaceborne imaging spectroscopy data.

Zusammenfassung

Die städtische Umwelt und ihre Prozesse wirken sich unmittelbar auf das menschliche Leben aus. Detaillierte und aktuelle Karten der städtischen Oberflächenmaterialien sind für Modellierer, die sich mit Meteorologie, Klimatologie und Ökologie beschäftigen, sowie für Behörden, die die Dynamik des städtischen Wachstums und die räumliche Entwicklung verstehen wollen, von großer Bedeutung. Die Kartierung der städtische Oberflächenmaterialien ist jedoch aufgrund der komplexen räumlichen Muster eine Herausforderung. Eine bewährte Quelle für aktuelle Informationen ist die Fernerkundung, wie die weit verbreitete Nutzung von SAR-, LiDAR- und optischen Daten zeigt. Daten von bildgebenden Spektrometern können hierbei durch die feine und kontinuierliche Abtastung des elektromagnetischen Spektrums detaillierte spektrale Merkmale von Oberflächenmaterialien erkennen, was mit multispektralen oder RGB-Bildern nicht mit der gleichen Genauigkeit erreicht werden kann. Bislang wurden in zahlreichen Studien zur Kartierung von städtischen Oberflächenmaterialien Daten von flugzeuggestützten abbildenden Spektrometern mit hoher räumlicher Auflösung verwendet, die ihr Potenzial unter Beweis stellen und gute Ergebnisse liefern. Im Vergleich zu diesen Sensoren haben weltraumgestützte abbildende Spektrometer eine regionale oder globale Abdeckung, eine hohe Wiederholbarkeit und vermeiden teure, zeit- und arbeitsaufwändige Flugkampagnen. Allerdings liegt die räumliche Auflösung der aktuellen weltraumgestützten abbildenden Spektroskopiedaten bei etwa 30 m, was zu einem Mischpixelproblem führt, welches mit herkömmlichen Kartierungsansätzen nur schwer zu bewältigen ist.

Das Hauptziel dieser Studie ist die Kartierung städtischer Materialien mit bildgebenden Spektroskopiedaten in verschiedenen Maßstäben und die gleichzeitige Nutzung des Informationsgehalts dieser Daten, um die chemischen und physikalischen Eigenschaften von Oberflächenmaterialien zu erfassen sowie das Mischpixelproblem zu berücksichtigen. Konkret zielt diese Arbeit darauf ab, (1) photovoltaische Solarmodule mit Hilfe von luftgestützten bildgebenden Spektroskopiedaten auf der Grundlage ihrer spektralen Merkmale zu kartieren; (2) die Robustheit der Stichprobe von städtischen Materialgradienten zu untersuchen; (3) die Übertragbarkeit von städtischen Materialgradienten auf andere Gebiete zu analysieren.

Zu diesem Zweck haben wir durch die Erstellung und Kombination von Spektralindizes (Ji et al., 2021a) Solaranlagen mit einer Gesamtgenauigkeit von rund 80% bis 90% erkannt. Diese Dissertation zeigt, dass der entwickelte Ansatz für eine genaue Erfassung von Photovoltaikanlagen geeignet ist. Wir haben auch gezeigt, dass das Konzept der Materialgradienten städtischer Oberflächen stichprobenrobust (Ji et al., 2020) und zwischen ähnlichen städtischen Gebieten übertragbar ist (Ji et al., 2021c). Mit diesen Ergebnissen kann das Konzept der urbanen Materialgradienten ein generischer Ansatz für die Stadtkartierung mit weltraumgestützten abbildenden Spektroskopiedaten sein. Die drei Teile dieser Dissertation verbessern somit die Anwendung von bildgebenden Spektroskopiedaten für die Erfassung städtischer Materialien, von einer klassischen merkmalsbasierten Methode zum neuen Konzept der städtischen Gradienten, von luftgestützten zu weltraumgestützten Daten, von der reinen Pixeldetektion zur Lösung des Problems der gemischten Pixel. Insbesondere durch die Einführung und Verbesserung des Gradientenkonzepts in der Stadtkartierung kann das Problem der gemischten Pixel angegangen werden, was ein vielversprechender Ansatz für die Analyse von abbildenden Spektroskopiedaten von aktuellen und zukünftigen weltraumgestützten Sensoren ist. Somit liefert diese Arbeit vielversprechende Ergebnisse für die Kartierung von städtischen Oberflächenmaterialien, indem sie einen auf physikalischen Merkmalen basierenden Ansatz vorschlägt und das generische Gradientenkonzept für städtische Materialstudien bestätigt und begründet. Weitere Arbeiten können auf diesen Ergebnissen aufbauen und könnten ein neues Feld für die Anwendung aktueller und zukünftiger weltraumgestützter abbildender Spektroskopiedaten eröffnen.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Urban environment and surface materials

1.1.1 Urban environment

Urbanization is one of the most pressing global challenges (Esch et al., 2013). About 55% of the world's population lives in urban areas, and this proportion is expected to rise to 68% by 2050 (DESA, 2018). Many countries face the challenge of balancing the needs of growing urban populations with environmental and societal systems (Lakes et al., 2021). In the long term, urbanization drives environmental change across multiple scales and in various ways, including local urban heat islands (Zhang et al., 2011; Schuster et al., 2017), biodiversity loss (McDonald et al., 2013), regional air pollution (Kennedy et al., 2009; Fenger, 1999), and global climate change (Wen et al., 2017). Therefore, a thorough understanding of the ecological processes shaping the urban environment is a key element for studying meteorology (Auer Jr, 1978), climatology (Seto and Shepherd, 2009), and ecology (Lakes and Kim, 2012) of cities. Sustainable development increasingly depends on the successful management of urban growth. Issues like cooling and shading potential, (Upreti et al., 2017), urban green distribution (Wolch et al., 2014), or the condition of vegetation in terms of drought and ecological quality (Grote et al., 2016) rely on solid urbanization management. A special case is that cities have the potential to turn the COVID-19 crisis into an opportunity nowadays. The current COVID-19 pandemic has lighted existing problems and inequalities in urban planning and management, e.g., cramped housing conditions (Nix et al., 2021), difficulty in maintaining a clean water supply (Feizizadeh et al., 2021), and inadequate sanitation facilities in informal settlements (Corburn et al., 2020). The need for public health interventions ranges from physical distancing, frequent hand washing, solid waste disposal,

provision of safe water and sanitation, which underscores the urgency of eradicating urban poverty and improving housing and infrastructure (Habitat, 2020; Austrian et al., 2020).

Therefore, continued urbanization drives the need to understand urban areas in order to improve and promote the environmental and human sustainability of cities worldwide (Weng and Quattrochi, 2018). Contemporary urbanization has the opportunity to foster ecological sustainability not only driven by increased environmental awareness, but also through returns from innovation, productivity, and efficiency (Grimm et al., 2008; Seto et al., 2010). To ensure that policy maintains the benefits of urbanization and anticipates or manages the negative consequences of urban growth, the interactions between socioeconomic and environmental processes in urban landscapes need to be better explored (Alberti, 2005; Lakes and Kim, 2012). To this end, a comprehensive understanding of urban materials is required.

1.1.2 Complexity of urban surface materials

Urban surface materials highly affect Earth's ecosystem processes in multiple ways. For example, the physical properties of urban surfaces influence urban micro-, meso-, and macro-climate, such as urban heat islands. Moreover, the albedo of surface materials impacts the radiative forcing of urban areas (Menon et al., 2010), leading to global warming. Surface runoff and impervious surfaces also affect hydrological processes and need to be considered simultaneously in urban planning to prevent flooding (Shao et al., 2019). Furthermore, urban surface materials, such as asbestos on the roof, also affect human health and well-being (Kantzioura et al., 2012; Yan et al., 2012). Urban surface material maps serve as fundamental information to better understand and plan for sustainable and habitable urban conditions. Therefore, knowing the details of urban surface materials is essential, e.g., specific heat capacity and emissivity of surface materials can be used to estimate heat fluxes and heat transfer, and their degree of imperviousness can be used to estimate water balance and surface runoff, etc.

However, mapping urban surface materials is difficult for several reasons. First, there are various natural and artificial surface materials in the urban environment. According to different classification hierarchies, urban surface materials can include several to dozens of classes. The abundances and classes of these materials are needed as inputs to urban climate models, but the materials are not easily distinguished. For example, red concrete and red clay can have very similar colors, but their daily and yearly temperature behaviors are very different, resulting in varying heat contributions to the urban environment.

Moreover, the complexity of mapping urban surfaces is also due to the heterogeneity of material distribution patterns. Urban surfaces generally consist of urban neighborhoods such as impervious structures (e.g., buildings and transportation networks), vegetation types

(e.g., parks, gardens, and agricultural fields), bare soil zones, and water bodies (Herold et al., 2002). In each general category, more varieties of urban neighborhoods are included. The heterogeneous mix of different materials in urban neighborhoods complicates the mapping of surface materials (Quattrochi and Ridd, 1994). In addition, heterogeneous regions dominated by small structures are characterized by the absence of pure materials, which further complicates the identification of surface materials (Roessner et al., 2001).

Furthermore, the high complexity of urban surface materials is also related to the specific local patterns of surface material compositions around the world. The occurrence of urban surface materials varies due to geographical conditions (Jilge, 2019), climate variations (Eliasson, 2000), political motivations (Pearsall and Pierce, 2010), historical reasons, etc., and therefore results in different combinations of urban surface materials in certain regions. For instance, residential areas of European cities commonly use roof tiles, while Northern American cities typically use slate as a roofing material for smaller single and multi-family houses (Herold et al., 2003). The diverse co-occurrence of urban surface materials poses greater challenges to the robustness and transferability of many urban surface mapping methods. Therefore, the complexity of urban surfaces hinders the detailed mapping of surface materials, which is of great importance for understanding urban dynamics in the face of rapid urbanization. In this scenario, Earth observation data can be used as a suitable solution to the mapping challenges.

1.2 Urban surface mapping using Earth observation (EO) data

Earth Observation (EO) is the collection of information about the physical, chemical, and biological systems of planet Earth. It monitors and assesses the state of and changes in the Earth's natural and man-made environment using technologies like remote sensing. Remote sensing provides a view from space and monitors widespread changes in the Earth sciences, particularly in areas such as meteorology, oceanology, hydrology, geology, geography, forestry, agriculture, and geodynamics (Kramer, 2012).

In the last decades, studies on urbanization emphasize the need for the collection and analysis of EO data (Wentz et al., 2014). The employment of EO data is of great benefit for objective and independent monitoring urban areas, and as a complement to traditional filed surveys or socioeconomic statistics, thus improving spatial resolution and interpretability (Potere et al., 2009; Palacios-Lopez et al., 2021). Recent researches have highlighted a number of approaches to land use data collection, characterization, and analysis that utilize

remote sensing images as source data to derive global data sets with high spatial resolution, including MODIS 500 urban land cover (Potere et al., 2009), Global Human Settlement Layer (GHSL) (Pesaresi et al., 2013; Corbane et al., 2019), GlobeLand 30 (Chen et al., 2015), GLOBELAND30-GLC309 (Chen and Chen, 2018), Global Urban Footprint (GUF) data set (Esch et al., 2017), and recent World Settlement Footprint (WSF) (Marconcini et al., 2020; Esch et al., 2022).

1.2.1 Mapping urban surfaces using optical EO data

Remote sensing (RS) is considered as one of the primary means of EO data collection, which measures electromagnetic radiation interacting with the atmosphere and objects on Earth (Zhu et al., 2018). The interactions of electromagnetic radiation with the Earth's surface can provide information not only about the distance between the sensor and the object, but also about the direction, intensity, wavelength, and polarization of electromagnetic radiation (Davis et al., 1978; Campbell and Wynne, 2011). Depending on the covered spectral range, remote sensing technology includes optical, thermal, and microwave detection systems (Small et al., 2018). Optical remote sensing is the detection of reflected solar radiation from objects in the spectral range of visible (VIS, 400 nm - 700 nm), near-infrared (NIR, 700 nm - 1300 nm), and shortwave infrared (SWIR, 1300 nm - 3000 nm). It is capable of discriminating physical and chemical properties of man-made materials and biophysical conditions of vegetation species (Lausch et al., 2016) and is therefore widely used for urban surface detection.

When monitoring complex urban surfaces using optical remote sensing data, three characteristics should be taken into account. First, the spatial resolution of the optical data affects urban surface detection. Spatial resolution refers to the size of a pixel of the image on the ground. A pixel is the smallest "dot" that makes up an optical image and essentially determines how detailed an image is. A pixel can be covered with only one surface material (called a pure pixel) or with various urban surface materials (called a mixed pixel). The reflectance of a pure pixel is defined as an endmember (Zare and Ho, 2013). Classical classification methods require endmembers extracted from pure pixels as input. However, despite other factors such as local surface complexity, in general, the lower the spatial resolution of the image is, the fewer pure pixels can be found. Therefore, the spatial resolution of the optical data directly affects urban surface mapping in terms of the level of detail and the applicability of the classification methods.

The second essential factor is the temporal resolution of the EO sensors. Since the urban surface material mapping is required as input for micro-, meso- and macro- climate models, capturing timely surface information and the temporal variation of surface materials

is essential for the analyses within these dynamic climate models. Although urban surface material maps do not need to be updated in a very short period of time compared to, for example, hazard maps that require daily or hourly updates, the annual production of surface material maps can be beneficial for climate model analyses.

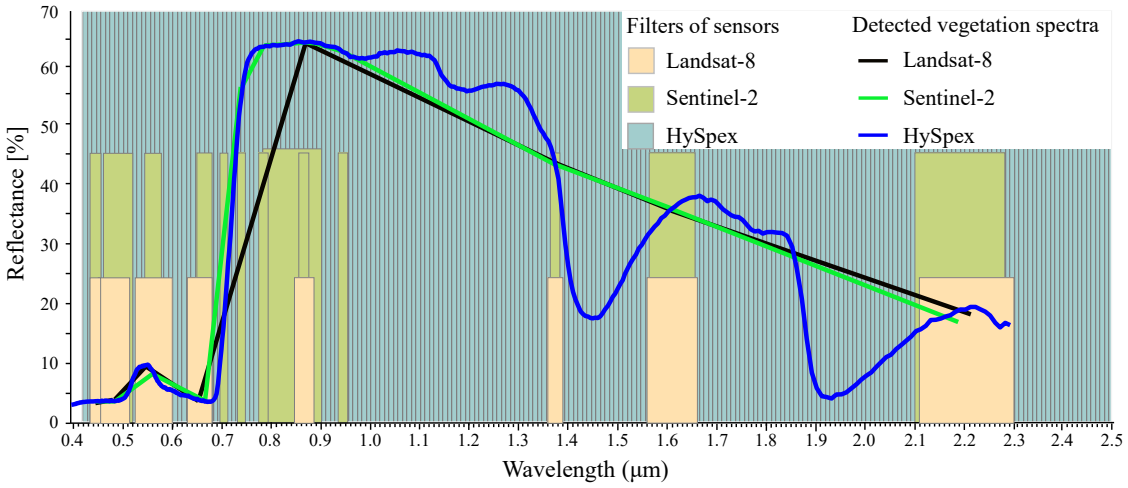


Fig. 1.1 The simulated vegetation spectra from Landsat, Sentinel-2, and HySpex.

In particular, the spectral resolution of EO sensors affects the detection of surface materials based on their diagnosed spectral features, thus compromising the efficiency and accuracy of urban surface mapping. For example, Fig. 1.1 shows simulated vegetation spectra acquired by Sentinel-2, Landsat, and HySpex with different spectral resolutions. In general, the visual spectral characteristics, the "red edge", and the rough structure of cells can be detected by all sensors. However, for the same vegetation, the spectra acquired by multispectral sensors (Landsat and Sentinel-2) are much coarser in spectral dimension than the imaging spectroscopy (IS) data acquired by HySpex. Vegetation biochemical information is missed by both multispectral sensors because they do not have the detailed spectral resolution in the 1400 nm to 2500 nm range. In addition, many urban surface materials are spectrally indistinguishable in multispectral data. Different urban surface materials can have similar colors and textures, so only very fine spectral resolution can identify their differences. For instance, slates, as a widely used roof material, show great confusion in classification with other materials such as asphalt, cobblestone, dark-shingle, and metal roofs (Franke et al., 2009). More spectra examples are shown in Fig. 1.2, the broader spectral range and fine spectral bands can greatly extend the identification capabilities of urban materials with similar colors and textures. In this sense, the fine and continuous spectral resolution of optical images is the essential factor for identifying and monitoring urban surface materials.

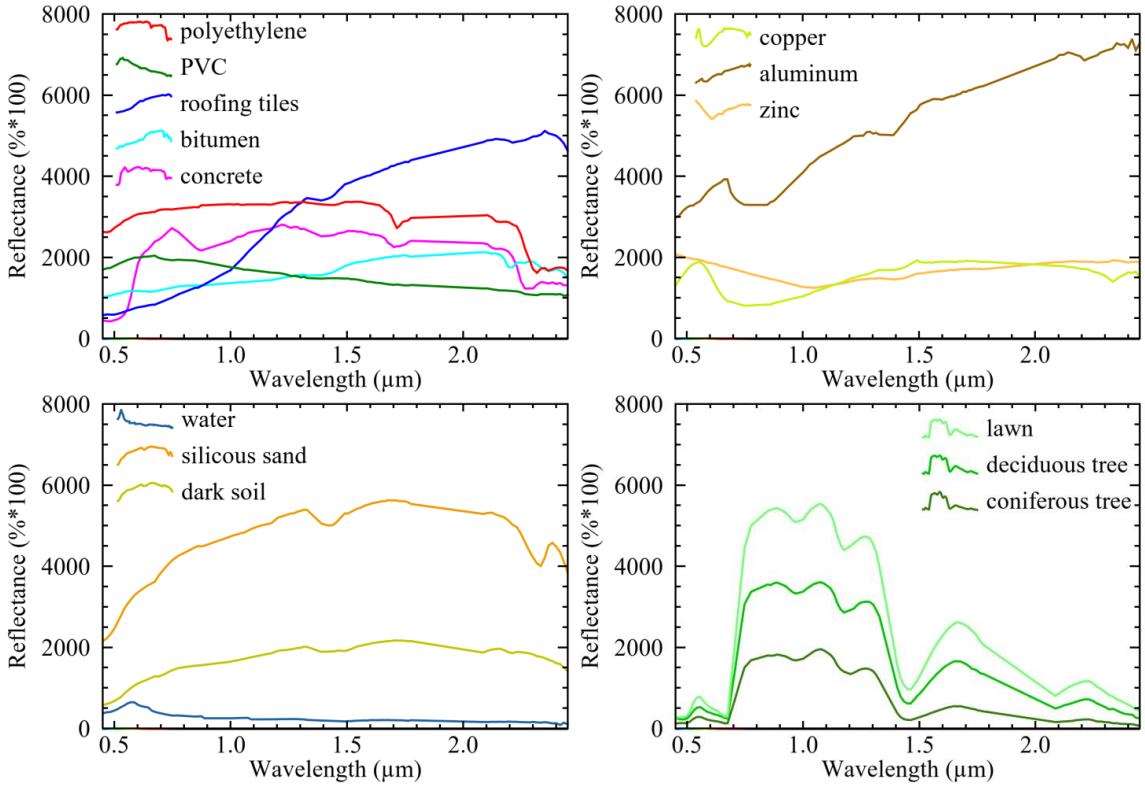


Fig. 1.2 The spectra collection of some urban surface materials collected from airborne HyMAP sensor (Heldens, 2010).

Spaceborne multispectral data have stable repeatability, relatively fine spatial and spectral resolution, as well as global coverage, and thus have been broadly used for urban land cover mapping (Kabisch et al., 2019). For example, Landsat sensors have been extensively employed for urban surface mapping since the 1970s. A typical method of urban mapping using Landsat data is spectral indices, including both vegetation indices (e.g. Normalized Difference Vegetation Index, NDVI) (Masek et al., 2000) and built-up indices (e.g. Normalized Difference Built-up Indicator, NDBI) (Zha et al., 2003), to seperate urban surfaces from non-urban lands. Furthermore, Ridd (Ridd, 1995) proposed viewing urban pixels as being composed of linear combinations of three generic land cover components, i.e. Vegetation, Impervious surface and Soil (VIS model). Moreover, Landsat data have also been used to study how urban surface has been changing for decades due to its long history and relatively high spatial resolution. The time series analysis using Landsat data has been further revolutionized the way of using Landsat data (Marconcini et al., 2020), and a majority of them were focused on change detection (Zhu, 2017). In addition, the Sentinel-2 mission launched in June 2015, with 13 spectral bands and up to 10 m spatial resolution, offers a great potential for addressing the satellite data requirements for finer scale mapping of human settlements.

Compared to Landsat sensors, Sentinel-2 offers higher spatial and spectral detail, as well as better quality of the geospatial information layer for describing urban surfaces (Pesaresi et al., 2016; Priem et al., 2019). However, these multispectral sensors are commonly used for urban land cover mapping, but lack of material differentiation capability due to coarse spectral resolution.

1.2.2 Surface materials identification using imaging spectroscopy (IS) data

Airborne IS data possess high spatial and spectral resolution, and can therefore provide details for urban surface mapping. For instance, the Digital Airborne Imaging Spectrometer (DAIS 7915) could acquire 72 spectral bands covering VIS, NIR, and SWIR, and produced a detailed surface material map including dark and red loose chippings, metal, asphalt, concrete, and many more (Segl et al., 2003; Roessner et al., 2001). Plaza et al. (2005) subsequently developed an Extended Morphological Transformation (EMT) method based on DAIS 7915 data to analyze spatial and spectral patterns simultaneously, and generated urban surface maps with more details. Recently developed airborne imaging spectrometers such as APEX, HyMAP, and HySpex (see Fig. 1.2), with hundreds of spectral bands, have extended the study of urban surface materials to a higher level (Franke et al., 2009; Heiden et al., 2007; Heldens, 2010; Priem et al., 2021). However, airborne IS data have been limited with relatively small coverage, expensive, time-consuming, and labor-intensive flight campaigns, so mapping of urban surface materials can only be done locally and lacks repeatability.

Spaceborne IS data overcome the limitations in data availability of airborne IS data and has considerable potential to become a promising data resource for the future. Several spaceborne sensors have been delivering or will deliver their archives, including the Japanese Hyperspectral Imager Suite (HISUI) (Matsunaga et al., 2019), Chinese HuanJing-1A (HJ-1A) (Lu et al., 2011), GaoFen-5 (GF-5) (Su et al., 2021), Italian PRecursore IperSpettrale della Missione Applicativa (PRISMA) (Cogliati et al., 2021), German DLR Earth Sensing Imaging Spectrometer (DESIS) (Alonso et al., 2019), and the upcoming German Environmental Mapping and Analysis Program (EnMAP) (Guanter et al., 2015; Bachmann et al., 2021), ESA's Copernicus Hyperspectral Imaging Mission for the Environment (CHIME) (Rast et al., 2021), and NASA's Surface Biology and Geology (SBG) (Cawse-Nicholson et al., 2021) missions. In particular, some sensors attempt to acquire and deliver data for specific targeted areas (e.g. DESIS and EnMAP), and in the near future sensors will seek global coverage of IS data (e.g. CHIME and SBG). In recent years, the development of these sensors has promoted numerous applications owing to the high spectral and temporal resolution and

regional or global coverage of these sensors. For example, some studies have attempted to use DESIS data for water quality mapping (Pinnel et al., 2021), automatic mapping of forests (Marshall et al., 2021), and detection of changes in solar panels (Ji et al., 2021b). The simulated EnMAP data have been widely used for mapping cities (Segl et al., 2012; Okujeni et al., 2015; Van der Linden et al., 2015; Jilge et al., 2019), forest (Cui et al., 2019; Cooper et al., 2021), and soil detection (Castaldi et al., 2019; Ward et al., 2020).

However, the relatively coarse spatial resolution of spaceborne IS complicates data implementation due to the lack of pure pixels or, in other words, full of mixed pixels (1.2.1 section). A number of studies have been working on tackling or bypassing this problem. For example, Yokoya et al. (2017) concentrated on the image segmentation between EnMAP and Sentinel-2 to improve the spatial resolution of the segmentation data. Okujeni et al. (2015) implemented an extended VIS model on the data and pointed out that the simulated EnMAP data can not fully overcome the spectral intraclass similarity and the spectral confusion caused by the presence of shaded regions. And Jilge et al. (2019) developed urban material gradients based on previous detailed material maps and generated fuzzy urban material gradient maps by applying these gradients to EnMAP data.

In summary, airborne IS data are capable of identifying detailed spectral signatures and providing spectrally pure pixels for surface materials classification, and most current studies on urban surface material mapping employed airborne IS data. However, they are limited by low spatial coverage, lack of availability, poor continuity, demanding workload, and high cost. Spaceborne IS data are considered to have promising potential for future urban surface material map because they have high repeatability and spectral resolution, but currently lack high spatial resolution, e.g., 30 m spatial resolution for DESIS, PRISMA, and EnMAP data. The resulting mixed pixel problem complicates the application of ongoing and upcoming spaceborne IS data for urban areas since classical classification methods (e.g. dimension reduction and regression methods) require spectrally pure pixels as input endmembers. Therefore, current airborne IS data are able to be used for urban surface material identification but are not a promising resource for a dynamic monitor of urban surface materials. In addition, the mixed pixel problem with applying spaceborne IS data should be solved to enable more efficient, accurate, and widespread application in the future.

1.3 State-of-the-art in urban materials mapping using IS data

In this section, the algorithms dealing with IS data are generally summarized and analyzed. Subsection 1.3.1 elaborates the state-of-the-art methods. In subsection 1.3.2, the spectral feature-based classification methods using IS data are thoroughly reviewed. Finally, subsection 1.3.3 presents studies on gradient mapping based on ordination methods using IS data.

1.3.1 General overview

This subsection outlines the state-of-the-art IS data classifiers according to the different spatial resolutions of data sources (e.g. airborne and spaceborne IS data), whether adding additional information (e.g. height information, thermal data, SAR data, or LiDAR data), classification unit (e.g. pixel-based, sub-pixel based, feature, or object-based classification unit), the pattern of mapping results (e.g. hard classification, soft classification, and gradient mapping), and a brief summary of their merits and weaknesses.

As aforementioned, the spatial resolution of airborne IS data is critical for urban surface classification, because most classifiers require many spectrally pure pixels as input. Airborne IS data commonly have a spatial resolution of 1 m to 10 m, making them well suited for finding spectrally pure pixels to classify urban surfaces (Roessner et al., 2001; Herold et al., 2003; Segl et al., 2003; Herold et al., 2004; Heiden et al., 2007, 2012; Degerickx et al., 2018; Jiang et al., 2019; Ji et al., 2021a). For instance, linear or nonlinear unmixing methods and endmember extraction are commonly applied with airborne IS data, which could provide spectrally pure pixels required by classifiers (Bioucas-Dias et al., 2012; Broadwater and Banerjee, 2009). When moving from airborne to spaceborne observations, a strong mixing of spectral signals occurs on most pixels. Therefore, few spectrally pure pixels can be extracted, making a hard classification at the urban material level unrealistic (Small, 2005).

Some classifiers seek better mapping accuracy by adding more information, which is also referred to as multi-modality IS data analysis. Image fusion of IS data and multispectral data can overcome the inconvenience of coarse spatial resolution of IS data. Due to sensor design considerations, the wealth of spectral information in IS data is often not complemented by extremely fine spatial resolution (Li et al., 2011). In particular, spaceborne IS data often have a spatial resolution of 30 m, which complicates their use for urban surface classification. Tremendous efforts have been made to fuse IS data with higher spatial resolution multispectral data (Yokoya et al., 2017; Xie et al., 2019; Dian et al., 2020). Moreover, since some buildings

and open spaces are covered with spectrally similar urban surface materials that hamper clear discrimination between them, adding additional height information could also improve the classification accuracy (Heiden et al., 2012). In addition, some studies incorporate temperature information acquired from thermal bands of remote sensing data (Segl et al., 2003; Eslami and Mohammadzadeh, 2015). These approaches could also enhance the exploitation of the information potential of the IS data. Furthermore, high density LiDAR data can also be involved and integrated as input data (Priem and Canters, 2016), in particular in the extraction of tree information (Heumann, 2011).

An alternative perspective of the urban mapping using IS data is the classification unit, ranging from sub-pixel-based and pixel-based to feature- or object-based. In sub-pixel classifiers, the spectral value of each pixel is assumed to be a linear or nonlinear combination of endmembers or pure pixels. Therefore, they are also referred to as spectral unmixing methods using linear or nonlinear models. Linear regression is widely used in practice because it is simple and generalizes well, while nonlinear regression is used for more complex nonlinear relationships and its solution is usually obtained by solving an approximate linear regression problem. Pixel-based classifier treats each pixel as the basic unit and then feeds pixels to the classifier as inputs. Many classical classification methods are based on this approach. For example, Franke et al. (2009) developed a pixel-based classifier that defines a hierarchy that uses spatial information from one level to constrain model selection at a higher complexity level. Then, three end-member selection procedures were used to identify the most representative end-members for each complexity level separately, which was proven to be particularly well suited for an urban environment. The object-based classifier is a segmentation technique that assigns a label to each pixel in the image so that pixels with the same label share certain visual features. In this case, classification is performed based on the objects and not on a single pixel (Liang et al., 2013; Makantasis et al., 2015).

Regarding the belonging of classified pixels to specific clusters, the classifiers of urban surface material mapping can be distinguished as hard classification, soft classification, and gradient mapping based on ordination. Hard classification, also called crisp or Boolean classification, is a common approach not only for urban surface mapping but also for mapping vegetation species and other objects. Soft classification, also referred to as fuzzy classification, considers the probability of pixels belonging to classes. Gradient mapping based on ordination attempts to describe surface materials or plant species composition as a continuum. Gradient mapping best preserves information of the original data and no prior pure pixel is required. In soft classification, the continuity of objects is preserved and little information is lost, but it is still based on a previous classification (Feilhauer et al., 2020).

Theoretically, a soft classification or gradient map can be converted into a hard classification map, but not vice versa.

1.3.2 Spectral feature-based material identification

Many studies have demonstrated that robust spectral features have the tremendous ability for urban surface material mapping (Kuehn et al., 2004; Keshava, 2004; Heiden et al., 2007). The absorption band positions, depths, and widths of spectra are correlated with diagnostic physicochemical material properties such as composition and abundance (Cloutis, 1996). In general, spectral feature-based detection using IS data has been developed primarily for minerals and vegetation, and is now being applied with spectral indices to other objects, including sea oil pollution detection and urban surface material identification(Kokaly, 2011; Ji et al., 2021a).

Researchers have found that IS data can provide greater mineralogical details, enabling a detailed geological, geochemical, and geothermometric survey of a target region (Ben-Dor and Kruse, 1995; Cloutis, 1996). Many studies have compared the ability of multispectral data and IS data in geological remote sensing and found that a smaller wavelength interval is irreplaceable in diagnosing soil or rock characteristics (Rowan et al., 1977; Hunt and Salisbury, 1978; Kruse, 1988; King and Clark, 1989; Clark et al., 1990). The spectral parameters of an absorption feature such as band minimum position, band depths, bandwidths, band areas, absolute reflectance, and ratios of these various parameters can be used to obtain compositional information (Cloutis, 1996; Heiden et al., 2007). The spectral parameters can be used to quantify, or at least rigorously constrain important physical and chemical properties, such as major, and in some cases minor, element chemistry, end-member abundances, and grain sizes (Cloutis, 1996).

In the last decades, more classification studies with IS data based on the spectral features of surface materials have been done. Heiden et al. (2007) implemented the interactive method to determine and evaluate diagnostic spectral features based on comprehensive field and airborne image spectral libraries of more than 21,000 spectra of surface materials widely distributed in German cities. The proposed feature functions describe the stand deviation, ratio, area, absorption depth and position, reflectance height and position, as well as the polynominal adaption 1st order of spectral features. Kuehn et al. (2004) proposed the Hydrocarbon Index (HI) based on the absorption of hydrocarbon-bearing materials at 1.73 μm and 2.31 μm with airborne IS data. The HI has been applied in the detection of oil spills (Kokaly et al., 2013; Leifer et al., 2012; Khanna et al., 2013; Garaba and Dierssen, 2018; Pelta et al., 2019; Liu et al., 2019), plastics (Levin et al., 2007; Garaba and Dierssen, 2018; Acuña-Ruz et al., 2018; Guo and Li, 2020), and other hydrocarbon-containing materials.

Keshava (2004) et. al applied band add-on (BAO) on the spectral angle mapper (SAM) classifier, and therefore increased the angular separability between two classes of spectra and improve the discrimination of very similar targets using only a fraction of the available spectral bands.

Therefore, material identification based on spectral features has been developed for a long time, and became a classical method. It has been applied to the identification of geological materials with even greater similarity and ease of confusion. Applying this technique to the identification of urban surface materials using IS data shows that it has great potential to deal with the complexity of surface materials. In particular, when the target material has identical spectral characteristics, spectral feature-based classification can be a straightforward and applicable method.

1.3.3 Gradient analysis based on ordination methods

Awareness has been growing among landscape ecologists that the gradient or continuous model can provide an accurate representation of landscape heterogeneity (Feilhauer et al., 2020). The gradient concept was introduced by Gleason (1926) for vegetation classification based on the vegetation species continuum, which was called floristic gradient. Subsequently, the floristic gradient was successfully applied to vegetation mapping using remote sensing data (Schmidtlein and Sassin, 2004; Schmidtlein et al., 2007; Feilhauer et al., 2011, 2014, 2020; Neumann et al., 2016; Neumann, 2017). Gu et al. (2015) began applying gradients in the urban environment to quantify tree species composition using IS data, but did not yet perform gradient analysis in urban surface material mapping.

Recently, Jilge et al. (2019) proposed to implement the gradient concept in urban surface material compositions, as it was assumed and tested that the co-occurrence of urban surface materials within urban neighborhoods are similar to the floristic gradient in vegetation species. The gradient concept based on the distribution patterns of urban materials was implemented differently than the spectral mixture analysis based on pixel information. Since the gradient technique assumes that all pixels are mixed and no spectrally pure pixels are required, it becomes a practicable solution for the spectral mixture problem. These material distribution patterns were summarized with a dimensionality reduction to describe the main gradual changes in material compositions, which are called urban surface material gradients. Then, the material gradients are extracted from a table that lists all material cover fractions for a training sample comprising a variety of material mixtures. A set of numerical scores is subsequently provided that indicates its position on the derived material gradients and which can be treated as a proxy of its material composition.

Since no prior determination of pure pixels is required, the urban material gradient could be a promising and general technique for mapping urban material with spaceborne IS data. However, there are many factors that need to be analyzed to determine whether this technique is transferable. These include, for example, the robustness of the sampling, the transferability to a specific area, the best method for dimensional reduction, and the uncertainty if there is a time difference between the surface material map and the HSI data.

1.4 Research objectives and structure of the thesis

1.4.1 Research objectives

As stated, urban surface material maps are an indispensable input to various urban applications, serving to measure the degree of urbanization or being directly related to urban environmental conditions. IS data is an applicable resource for the identification of discrete urban surface materials or for detecting the co-occurrences of specific surface materials. More research is needed to capture the diversity of surface materials and their spectral behavior in spaceborne IS data so that more operational and standardized image analysis techniques can be further developed. This would pave the ground for indicators that inform the physical and chemical complexity of materials in urban areas and their impact on climatic processes. Therefore, given the first proof of concept for the existence of urban gradients (Jilge et al., 2019), this thesis further explores urban complexity measured with imaging spectrometers and the general applicability of the gradient concept as a potentially globally applicable technique. Since the urban material gradient is a data-intensive method, where might the extensive training data come from? Since the gradient concept is a data-driven technique, what impact does the sampling strategy have on the gradient results? And can we transfer the developed gradient models to other areas? In this regard, three topics were defined within the scope of this thesis:

Objective 1: Detection of solar photovoltaic module using laboratory and airborne IS data.

The first objective focuses on the spatial-spectral complexity of imaging spectroscopy measurements of urban areas across the lab, field, and airborne scale in order to develop physical-based measures for the robust detection of a target material. Solar photovoltaic (PV) modules can absorb and transfer solar energy and contribute to reducing greenhouse gas emissions and thereby mitigating climate change. Due to its variety of available types and implementation on rooftops and at large fields, EO assessments of solar PV modules are characterized by a high spatial and spectral complexity. Information about the occurrence

and the condition of solar PV is important for statistical purposes, management as well as regional planning of sustainable energy consumption. Many studies have been focused on the interpretation and detection of solar PV with color aerial image and visual interpretation. Benefiting from the fine spectral resolution of IS data, it potentially becomes a promising resource to detect PV modules based on the physics-based spectral characteristics of PV. However, the enormous spectral variability makes it difficult to simply detect PV with IS data, including spectral inter-class similarity and intra-class variability. PV module consists of several material layers, commonly including EVA coatings. EVA, as a typical hydrocarbon-bearing material, is spectrally similar to polyethylene-covered open surfaces, roofing polyethylene, and synthetic turf on sports fields, which could cause the spectral inter-class similarity. PV modules include different PV types and are characterized by a high intra-class variability due to the different illumination and observation angles using EO sensors. In PV detection, the spectral variability caused by different tilt angles of PV or detection angles of sensors is common and has therefore attracted our attention. In order to compensate for this, IS data acquired in the lab, field, and airborne should be considered to detect spectral features that are robust across all scales and can lead to a widely applicable method for PV identification. In this context, the first objective of this thesis is to solve the inter-class variety and intra-class similarity related to solar PV modules and, in a broader context, to deal with the complexity of urban surface materials using IS data.

Knowledge about the spatial-spectral complexity of urban areas and the ability of imaging spectrometry techniques to decode it for the accurate detection of a target material is a prerequisite to understanding and exploring spaceborne IS data for surface material composition mapping in the next chapter.

Objective 2: Analysis of the sampling robustness in gradient mapping of urban material mixtures

Since IS data are proven to be more powerful in capturing the complexity of urban surfaces, utilizing these data over larger coverage, with temporal replicates would be more valuable for the dynamic climate models that study interactions within urban ecosystems. Ongoing and upcoming spaceborne IS data can meet these needs but can lead to mixed pixel problems due to the relatively coarse spatial resolution. The lack of spectrally pure pixels for training classifiers poses a challenge. However, the distribution of urban surface materials is not arbitrary. Theoretically, Tobler (1970) found that nearby things are more related than distant ones. In reality, characteristic patterns of urban neighborhoods occur very frequently; for example, industrial areas are often characterized by a co-occurrence of concrete, asphalt, and metal roofing; residential areas often feature roofing tiles, trees, and lawns. EO sensors would then measure typical mixtures of co-occurring urban surface materials for certain

neighborhoods. In this context, Jilge et al. (2019) proposed the concept of urban gradient and proved the existence of urban surface material gradients. As a promising technology, surface urban material gradients could be a generic way to apply urban material mapping using spaceborne IS data. However, the stability of this technique has yet to be evaluated as it is determined in a data-driven manner. Therefore, the second objective of the dissertation is to analyze the sampling robustness of the urban surface material gradients and to find or evaluate a robust way for urban surface material mapping in a broader context.

Objective 3: Analysis of area transferability of urban material gradients

Gradient analysis appears to be a promising approach to address the problem of mixed pixels occurring with spaceborne IS data. Since gradients are generally determined in a data-driven manner, they may be only suitable locally, requiring additional field data collection if the gradients are transferred to other unknown areas. Such data collection can be an expensive and time-consuming task. After conducting the sampling robustness analysis of urban material gradients, the influence of such minor movements of sampling strategies was found to be marginal. In view of a broader application of the urban gradient technique, the evaluation of its transferability is the next necessary step. Thus, the third objective of this thesis is to analyze the area transferability of urban material gradients. Specifically, two questions are raised: (i) Are urban material gradients transferable between two study sites? (ii) What influences the transferability of urban material gradients? Although other factors, such as sample size and distance between samples, as well as ordination method, could also affect the physical significance of urban gradients, the most pressing questions are addressed in this thesis. In this regard, gradient analysis seems to be an applicable solution for generic mapping large coverage of urban surface materials to cope with the complexity of urban surface materials.

1.4.2 Structure of the thesis

The thesis is structured as follows. Chapter 1 provides a general introduction to the topic of urban material mapping using IS data. In addition, the motivation, as well as the objectives of this dissertation, are outlined. In Chapter 2 to Chapter 4, the core research of the cumulative thesis is included in terms of three stand-alone manuscripts that have been published in international, peer-reviewed journals. Chapter 5 summarizes the main findings and the contribution of the dissertation.

Chapter 2	*Solar photovoltaic module detection using laboratory and airborne imaging spectroscopy data.* Published in Remote Sensing of Environment (2021), 266, 112692.
Chapter 3	*Sampling robustness in gradient analysis of urban material mixtures.* IEEE Transactions on Geoscience and Remote Sensing (2020).
Chapter 4	*Are urban material gradients transferable between areas?* International Journal of Applied Earth Observation and Geoinformation (2021), 100, 102332.

1.4.3 Authors' contributions to the individual chapters

Chapter 2 Wieke Heldens, Uta Heiden and Marion Schroedter-Homscheidt conceived the idea. Susanne Weyand and Maron Schroedter Homscheidt organized the airborne campaign and collected validation data at Oldenburg test sites. Uta Heiden and Andreas Hueni organized the laboratory goniometer measurements and assisted in the post-processing and interpretation. Wieke Heldens and Uta Heiden provided the HyMap spectral library. Chaonan Ji performed the algorithm development with the support of Wieke Heldens, Martin Bachmann, and Uta Heiden. Martin Bachmann and Julian Zeidler support the technical parts of the study. Tobia Lakes, Hannes Feilhauer, Wieke Heldens, Uta Heiden, Marion Schroedter-Homscheidt and Martin Bachmann supervised the study. Susanne Weyand, Annekatrin Metz-Marconcini, and Marion Schroedter-Homscheidt clarified the background of the study. Chaonan Ji and Susanne Weyand conducted the original draft preparation of the manuscript. All authors contributed to the final review and editing of the manuscript.

Chapter 3 Chaonan Ji contributed to the literature review, development of the methods, writing the manuscript, planning the experiments, data collection, data analysis, programming, and argumentation. Marianne Jilge contributed to the development of the methods, review of the manuscript, planning of the experiments, data collection, and visualization. Uta Heiden assisted with data collection, review of the manuscript, review of results, argumentation, and supervision. Marion Stellmes contributed to the development of the methods, data analysis, and argumentation. Hannes Feilhauer provided assistance with development of the methods, review of the manuscript, design of experiments, data analysis, review of results, programming, and supervision.

Chapter 4 Chaonan Ji contributed to coding, writing, and funding. Uta Heiden assisted with data curation, design, supervision, and writing. Tobia Lakes contributed with visualization, writing, and supervision. Hannes Feilhauer provided support with conceptualization, methodology, supervision, and writing.

Chapter 2

Solar photovoltaic module detection using laboratory and airborne hyperspectral data

Chapter 2

Solar photovoltaic module detection using laboratory and airborne imaging spectroscopy data

Remote Sensing of Environment 266 (2021) 112692

Chaonan Ji, Martin Bachmann, Thomas Esch, Hannes Feilhauer, Uta Heiden, Wieke Heldens, Andreas Hueni, Tobia Lakes, Annekatrin Metz-Marconcini, Marion Schroedter-Homscheidt, Susanne Weyand, and Julian Zeidler

https://doi.org/10.1016/j.rse.2021.112692

Received 4 May 2021; Received in revised form 16 August 2021; Accepted 2 September 2021; Available online 16 September 2021

Abstract

Over the past decades, solar panels have been widely used to harvest solar energy owing to the decreased cost of silicon-based photovoltaic (PV) modules, and therefore it is essential to remotely map and monitor the presence of solar PV modules. Many studies have explored on PV module detection based on color aerial photography and manual photo interpretation. Imaging spectroscopy data are capable of providing detailed spectral information to identify the spectral features of PV, and thus potentially become a promising resource for automated and operational PV detection. However, PV detection with imaging spectroscopy data must cope with the vast spectral diversity of surface materials, which is commonly divided into spectral intra-class variability and inter-class similarity. We have developed an approach to detect PV modules based on their physical absorption and reflection characteristics using airborne imaging spectroscopy data. A large database was implemented for training and validating the approach, including spectra-goniometric measurements of PV modules and other materials, a HyMap image spectral library containing 31 materials with 5627 spectra, and HySpex imaging spectroscopy data sets covering Oldenburg, Germany. By normalizing the widely used Hydrocarbon Index (HI), we solved the intra-class variability caused by different detection angles, and validated it against the spectra-goniometric measurements. Knowing that PV modules are composed of materials with different transparencies, we used a group of spectral indices and investigated their interdependencies for PV detection with implementing the image spectral library. Finally, six well-trained spectral indices were applied to HySpex data acquired in Oldenburg, Germany, yielding an overall PV map. Four subsets were selected for validation and achieved overall accuracies, producer's accuracies and user's accuracies, respectively. This physics-based approach was validated against a large database collected from multiple platforms (laboratory measurements, airborne imaging spectroscopy data), thus providing a robust, transferable and applicable way to detect PV modules using imaging spectroscopy data. We aim to create greater awareness of the potential importance and applicability of airborne and spaceborne imaging spectroscopy data for PV modules identification.

2.1 Introduction

Due to the increasing energy demand (Wolfram et al., 2012; Sorrell, 2015), the need of cutting down greenhouse gas emissions (Zhang et al., 2019) and the ongoing energy transition process with substantial subsidies (Markard, 2018), the number of solar photovoltaic (PV) modules in operation has increased rapidly in recent years (Tao and Yu, 2015; Green, 2019). Several stakeholders such as environmental authorities, grid operators, manufacturing industries or energy system modelers are interested in monitoring PV system locations and areas, but accurate and publicly accessible databases are not available. Furthermore, these databases need continuous and regular updates. Although in-situ data can be collected through field surveys or citizen science projects, they are costly and/or time-consuming.

In this context, Earth Observation (EO) data offer a suitable alternative. EO data can provide the necessary spatial and temporal resolution to monitor PV modules on a large scale. Promising results have been achieved using color aerial imagery (Malof et al., 2016a; Yu et al., 2018; de Hoog et al., 2020). Malof et al. (2016a) investigated an approach based on supervised random forest classification to automatically identify distributed PV arrays using color aerial images with a spatial resolution of 0.3 m $\times$ 0.3 m, and achieved 72% precision and 80% recall. Yu et al. (2018) developed DeepSolar, a deep learning framework that analyzed color spaceborne imagery with a spatial resolution of 0.3 m $\times$ 0.3 m to identify the locations and sizes of solar PV modules. The resulting precision was 93.1% and recall was 88.5% in residential areas, while precision was 93.7% and recall 90.5% in non-residential areas. Leveraging its high accuracy and scalability, they constructed a comprehensive high-fidelity solar deployment database for the US. However, identifying solar PV modules across large regions remains challenging due to the requirement of high-resolution (typically 0.3 m/pixel or finer) imagery, difficult identification of solar PV modules in many situations (such as dark PV modules on dark roofs), and confusion of many other types of structures (such as solar hot water systems, roads, and even pools) to PV modules (de Hoog et al., 2020). This is because PV modules are composed of materials that typically include fully transparent glass covers for protection, highly transparent Ethylene Vinyl Acetate (EVA) films, and the core PV cell. In addition to these reasons, these methods require large, elaborated and pixel-accurate labeled data sets for training and validation (Malof et al., 2016a,b; Yuan et al., 2016; Camilo et al., 2018).

Instead of providing only RGB broad band spectra as color aerial imagery, imaging spectroscopy data can generally improve the separability of surface materials since its near continuous spectral information, with hundreds of narrow spectral bands can map the material-specific absorption characteristics (Herold et al., 2004; Heiden et al., 2007). Thus, more detailed spectral properties of PV modules can be derived from imaging spectroscopy

data. So far, very few studies focus on PV detection with imaging spectroscopy data. Czirjak et.al (Czirjak, 2017) showed that PV modules have a unique spectral signature that is consistent across multiple manufacturers and construction methods and is therefore detectable in imaging spectroscopy data, i.e., using adaptive cosine estimator to detect PV modules. In addition, Czirjak et.al (Czirjak, 2017) developed the Normalized Solar Panel Index (NSPI) to mitigate false positives by eliminating pixels that do not exhibit key spectral features of the reflectance spectrum of PV panels. The NSPI is designed to detect the steep increase in reflectance that typically occurs in spectral signatures of solar PV modules around 1.00 μm. Karoui et al. (2019) attempted to use Non-negative Matrix Factorization (NMF) algorithms to apply Linear Spectral Unmixing (LSU) on imaging spectroscopy data for PV detection. Their study concludes that the proposed approaches (Grd-Part-NMF and Multi-Part-NMF) are superior to the previous ones (Grd-NMF and Multi-NMF), which is a promising progress. However, it is important to note that the previous NMF approaches (Grd-NMF and Multi-NMF) are not PV detection approaches. Moreover, only one mean spectrum of the ground-measured PV modules spectra was considered as the known spectrum but different types of PV were not considered, which means a lack of variation of the PV spectra in the training phase. In addition, Karoui et al. (2019) did not consider the spectral variability caused by different inclination or detection angles, which is a limitation of the linear unmixing methods in principle since these methods are generally considered when the landscape of the observed scene is flat and the irradiance is homogeneous (Dobigeon et al., 2013). Furthermore, Karoui et al. (2019) did not consider materials that have similar spectra to PV panels, such as polyethylene materials and oil, which have similar double absorption feature at 1.73 μm due to their hydrocarbon content, and water, which has similar low reflectance in the VNIR region. For these reasons, Karoui et al. (2019) attempted the creative and meaningful experiment on PV detection using imaging spectroscopy data with LSU, but the detection of PV modules were not accurate enough, while a simple one-class classification generally achieved better results.

Therefore, PV modules detection using imaging spectroscopy data should focus on the physical characteristics and the spectral uniqueness of PV modules. PV modules commonly consist of several layers, including fully transparent glass covers for protection, highly transparent EVA films, and the core PV cell. EVA is a hydrocarbon-bearing material, so regardless of how well EVA transmits solar energy, the hydrocarbon absorption exists at 1.73 μm. Crystalline silicon (C-Si), as a common PV cell material, has a strong absorption in the visible (VIS) region, resulting in low reflectance of PV modules in the VIS region. In addition, it has a decreasing absorption between 0.99 μm and 1.15 μm, resulting in a steep reflectance increase in this spectral region. Moreover, like most hydrocarbon surface

materials, PV modules have a strong absorption around 2.2 μm. However, PV detection using imaging spectroscopy data must cope with the vast spectral diversity of urban materials and related characteristics, commonly classified as intra-class variability and inter-class similarity. Intra-class variability means the spectral variability within the material class, and inter-class refers to the spectral similarity among different material classes (Zhang et al., 2006; Somers et al., 2011). Intra-class variability can be caused by several factors, such as color, coating, degradation of the material and illumination of the material as well as preprocessing of the acquisition data (Heiden et al., 2007). In PV detection, the spectral variability caused by different tilt angles of PV or detection angles of sensors is common and has therefore attracted our attention. In addition, polyethylene covered open surfaces, roofing polyethylene and synthetic turf on sports fields, which are hydrocarbon-bearing materials similar to EVA in PV modules, could cause the spectral inter-class similarity, and therefore are another problem to be addressed in PV modules detection.

The objective of this study is to detect PV modules using airborne imaging spectroscopy data. Specifically, we aim to address 1) the spectral intra-class variability caused by different viewing and illumination angles, which is always present in PV detection; 2) the spectral inter-class similarity that occurs mainly between PV modules and other hydrocarbon-bearing materials; 3) as well as to apply and validate the developed spectral indices on the city of Oldenburg, Germany. To address these questions, we firstly identify specific spectral features of PV in the optical spectral range and introduce spectral indices based on laboratory spectra-goniometric measurements with different detection angles and a large labeled HyMAP image spectral library. These indices are then applied to airborne HySpex images acquired over Oldenburg, Germany.

2.2 Data and study area

2.2.1 Laboratory spectra-goniometric spectral library

Five materials were measured with the ASD spectrometer with a 3° field of view installed on the LAGOS goniometer (Schopfer et al., 2008) (Fig. 2.1A), including two bitumen materials for roof covers, a monocrystalline PV module, a polycrystalline PV module, and a hydrogen carbonate (PVC) material normally applied on large flat roofs (Table 2.1). For each material, a total of 61 measurements with different detection angles were collected (Fig. 2.1B) as well as one measurement with the white reference. Specifically, the 61 measurements cover zenith angles of 0° to 75° with an interval of 15° and the azimuth angles of 0° to 330° with an interval of 30°. These measurements were labeled by alphabet for zenith angle (from

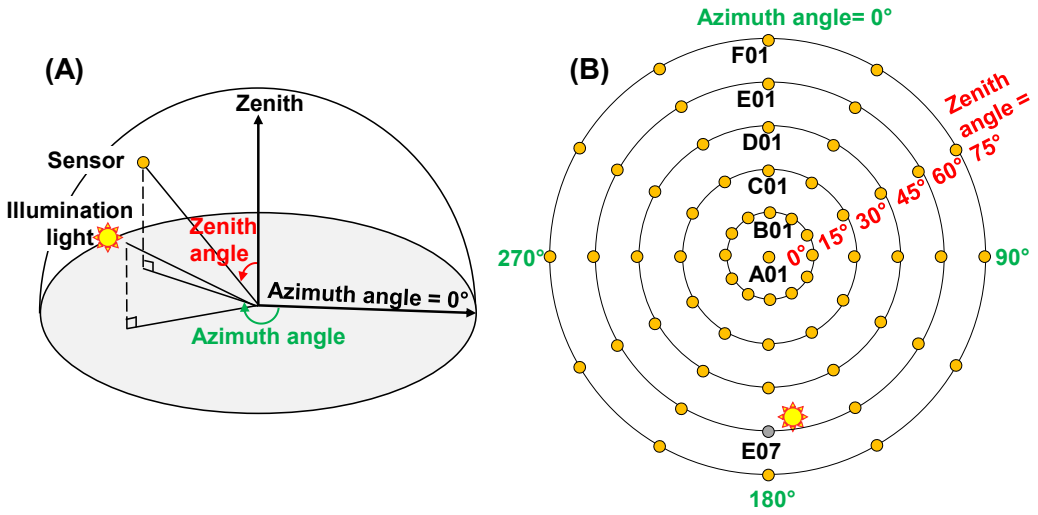


Fig. 2.1 The measurement set-up with an ASD sensor in a goniometer. (A) The material was placed on the gray platform and measured with fixed illumination light and a movable ASD sensor. (B) the arrangement of the total 61 measurement positions and their labels. Detection E07 was treated as an abandoned measurement because it was close to illumination light and affected.

A to F) and by number for azimuth angle (from 1 to 12) except for zenith angle = 0°. The measurement at E07 (zenith angle is 60°, and azimuth angle is 180°) was excluded, because the illumination source (55° of zenith angle, and 171° of azimuth angle) was between the sample and the optic of the spectrometer. Raw data were recorded as radiance and processed to reflectance by normalizing the radiance with the white reference. Therefore, a total of 60 reflectance spectra were available for each material. Each spectrum covers the spectral range from 350 to 2500 *nm*, with a spectral resolution of 1 *nm*.

2.2.2 HyMap image spectral library

An image spectral library extracted from airborne hyperspectral data was also included in this study. The library was mainly derived from imaging spectroscopy data recorded over Munich, Dresden, Potsdam, and Berlin, Germany in 1999, 2000, 2004, and 2007 (Segl et al., 2003; Heiden, 2004; Bochow, 2010; Heldens, 2010). All data were acquired with the HyMap sensor (Cocks et al., 1998). The radiometrically and atmospherically corrected HyMap data have 128 bands, of which three bad bands were removed, which were the first band of the visible (VIS) spectrometer, and the first and second band of the near-infrared (NIR) spectrometer (Heldens, 2010). The spectral library was extracted based on the method of Segl et al. (2003), developed by Heiden et al. (2012) and extended by Heldens (2010). It

Table 2.1 Materials and spectra in the laboratory spectra-goniometric measurements and HyMap image spectral library.

Laboratory spectra-goniometric measurements			
Materials	Spectra number	Color	Detail
Bitumen material A	60	red	age∼2017
Bitumen material B	60	gray	age∼2017
PV material A	60	black	monocrystalline PV cell
PV material B	60	dark blue	polycrystalline PV cell
PVC	60	black	hydrogen carbonate

HyMap image spectral library

Materials	Spectra	Materials	Spectra	Materials	Spectra
Roofing tiles	624	Roofing tar	15	Siliceous sand	31
Roofing concrete	352	Roofing glass	44	Humus soil	96
Aluminum	188	Vegetated roof	108	River	466
Copper	123	Concrete	157	Pond	183
Zinc	159	Asphalt	339	Pool	34
Polyvinyl chloride (PVC)	244	Concrete pavement	10	Coniferous trees	248
Roofing polyethylene	359	Cobblestone	10	Deciduous trees	277
Polyethylene surface	89	Loose chippings	184	Dry vegetation	19
Tartan	22	Railway tracks	65	Meadow	187
Synthetic turf	264	Vegetated railway tracks	28	Lawn	415
Roofing bitumen	287				

contains 5627 labeled spectra of 31 material classes. Each spectrum in this library has 125 spectral bands ranging from 450 to 2500 *nm*. The HyMap spectral library includes several polyethylene materials, i.e., roofing polyethylene, polyethylene surface, and synthetic turf, which have a hydrocarbon absorption similar to PV modules. This allowed us to collect the spectral features of these similar materials, and remove them in the PV detection.

2.2.3 HySpex images

Ten imaging spectroscopy data sets were collected from a flight campaign carried out in July 2018 covering Oldenburg with the HySpex sensor (see Fig. 2.2). The HySpex system has two cameras covering the spectral ranges of visible near-infrared (VNIR) and short-wave infrared (SWIR) region. The VNIR sensor records the spectral range from 416 to 992 *nm* with 160 channels at a spatial resolution of 0.6 m $\times$ 0.6 m. The SWIR sensor covers the spectral range from 968 to 2498 *nm* in 256 channels at a spatial resolution of 1.2 m $\times$ 1.2 m. To work with the same reference frame and the whole spectral range, the VNIR and SWIR images were co-registered (Schwind et al., 2014) and resampled to the same spatial resolution of the SWIR sensor, which is 1.2 m $\times$ 1.2 m. More details about the characteristics of the HySpex system are provided in Köhler (2016).

HySpex Level 2A data were provided for this campaign. After system correction, the data were ortho-rectified, and the surface reflectance was calculated with the ATCOR4 atmospheric correction software (Richter et al., 2011) for each HySpex flight line. This pre-processing was carried out by OpAIRS of the Remote Sensing Technology Institute (IMF) of German Aerospace Center (DLR), and described in detail in Köhler (2016). It should be noted that the uncertainty in the relative geolocation between two adjacent flight lines can be up to 3 pixels, and that a further systematic/non-systematic displacement to the validation data of 1-2 pixels exists. The effects are described in Section 2.3.4.

2.2.4 Study area

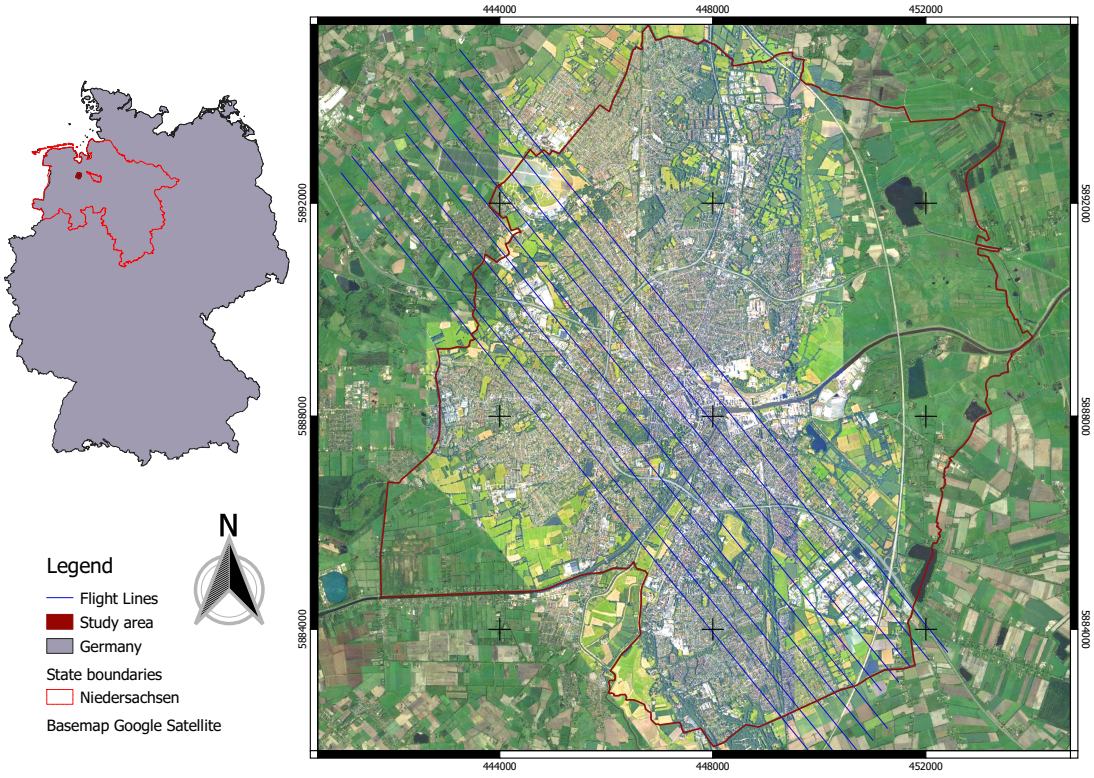


Fig. 2.2 The study area - Oldenburg in the northwest of Germany, covered with ten HySpex images.

The city of Oldenburg is located in the northwest of Germany (see Fig. 2.2), and covers an area of 103 km^2. The study area captured by the HySpex images includes a variety of building types with different installations of PV modules. A large PV power plant is located on the old airfield in northwestern Oldenburg. Slightly further south are two university campuses, Haarentor and Wechloy, and the Institute for Networked Energy Systems of the

DLR. The city center of Oldenburg is dominated by dense perimeter block developments with varying roof materials and few open spaces. The south of Oldenburg is characterized by several industrial areas with halls and warehouses, as well as large areas with many semi-detached and detached houses, some of which are covered with PV modules.

2.3 Methods

To capture and describe the spectral characteristics of PV modules, we applied a set of spectral indices using imaging spectroscopy data. The study workflow is shown in Fig. 2.3.

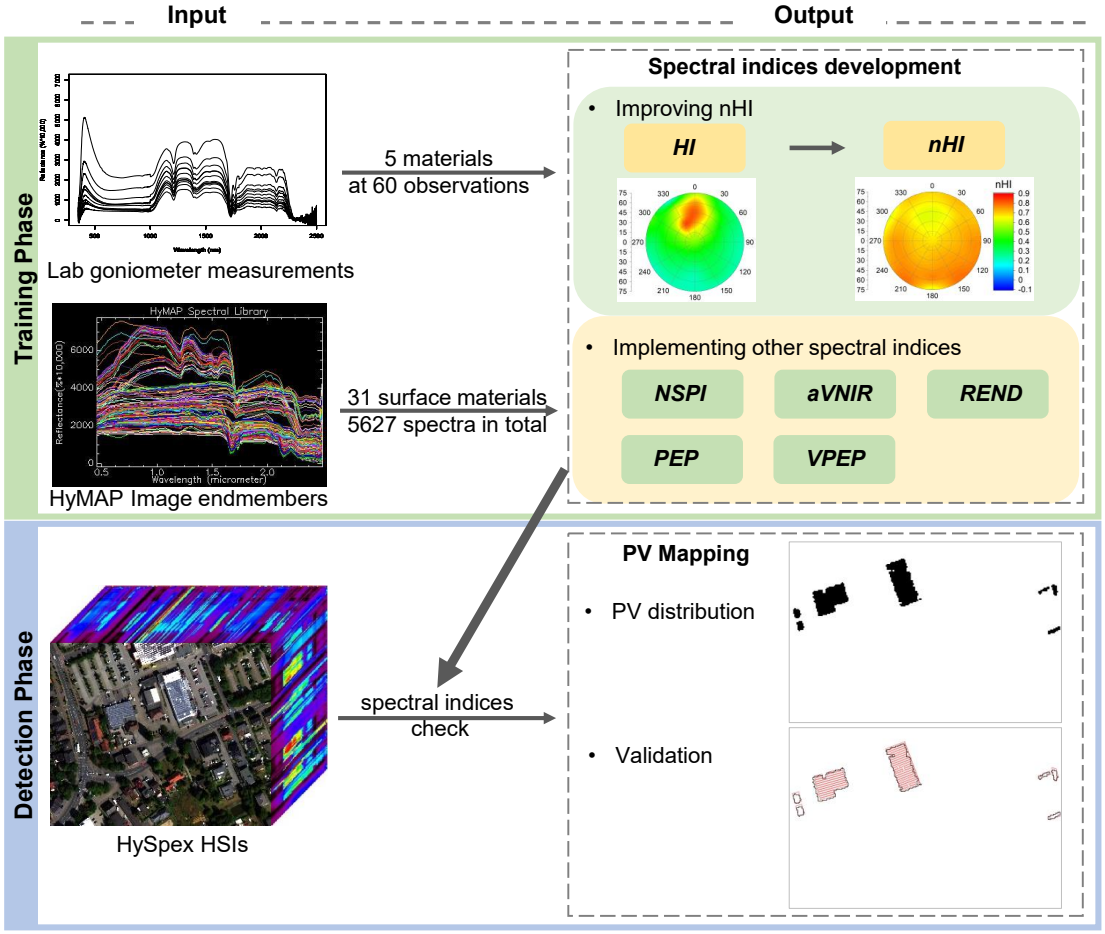


Fig. 2.3 The study workflow.

2.3.1 Hydrocarbon index normalization

The diagnostic spectral characteristics of hydrocarbon in the SWIR were reported by Cloutis (1989), which revealed the hydrocarbon absorption feature centered near 1.73 μm arising

from the various C-H stretching overtones and combination bands. Since then, the basic ability of hyperspectral systems was explored to detect hydrocarbon features in the SWIR. Hörig et al. (2001) realized this capability using airborne HyMap imaging spectroscopy data to map and delineate oil-contaminated soils based on the absorption feature. Based on this, Kuehn et al. (2004) proposed the Hydrocarbon Index (HI) (see Fig. 2.4 and Equation 2.1) that measures the depth of the spectral absorption at 1.73 μm to identify the presence of hydrocarbon-bearing material. This HI converts multi-band data into a single band, which is straightforward to use for detecting the presence of hydrocarbon-bearing materials.

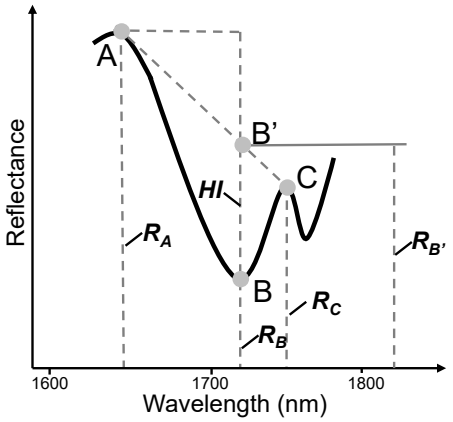


Fig. 2.4 Demonstration of the spectral absorption in the reflectance of hydrocarbon-bearing materials (modified from Kuehn et al. (2004)). A, B, and C are the points on the reflectance line, and B' is a point on the continuum line of points A and C. Accordingly, their reflectance values are R_A, R_B and R_C, and wavelengths are λ_A, λ_B and λ_C. The distance between points B and B' were defined as HI (Kuehn et al., 2004).

$$HI = R_{B'} - R_B \tag{2.1}$$

where

$$R_{B'} = (\lambda_B - \lambda_A)\frac{R_C - R_A}{\lambda_C - \lambda_A} + R_A \tag{2.2}$$

However, HI does not consider the spectral variation of hydrocarbon-bearing materials, i.e., the spectral intra-class variability of PV modules due to color, coating, degradation of the material and orientation of the material to the sensor etc. (Heiden et al., 2007; Clark and Roush, 1984; Clark et al., 2003; Sahib, 2019). Clark and Roush (1984) proposed a consistent band depth concept to reduce topographic and atmospheric effects by calculating band depth with the support of spectral continuum. By combining the concepts of band depth and HI, we performed a normalization procedure for the HI to minimize the influence of different

detection angles. The normalized HI, called nHI, is calculated by dividing the HI by the $R_{B'}$ (see Equation 2.3).

$$nHI = \frac{HI}{R_{B'}} = \frac{R_{B'} - R_B}{R_{B'}} \tag{2.3}$$

The HI was calculated by using 1705 nm, 1729 nm and 1741 nm as points A, B, and C in Kuehn et al. (2004). Since nHI uses the concept of continuum-removed absorption band and is calculated by dividing the band depth of each channel by the reflectance at the band center, points A and C should be at the spectral shoulders. Therefore, we selected 1669 nm and 1746 nm as the points A and C. The point B is still 1728 nm, which is the center of the hydrocarbon absorption feature. Given the different spectral resolutions of sensors used in this study, the exact wavelengths of points A, B, and C can be slightly modified and adjusted to suit particular sensors.

2.3.2 Additional spectral indices

Instances of different material classes may exhibit highly similar spectral features, which is referred to as inter-class similarity. The hydrocarbon absorption feature of PV modules at 1.73 μm is not unique, while other hydrocarbon-bearing materials also exhibit this feature. Therefore, these polyethylene materials such as roofing polyethylene, polyethylene surface, and synthetic turf, should be constrained by additional specific indices in case they are misclassified as PV modules. Thus, four additional spectral indices are developed in this study to accurately distinguish PV modules from other hydrocarbon-bearing materials. Including the above-mentioned nHI, a total of six spectral indices based on the physical characteristics and corresponding spectral features of PV modules are shown in Fig. 2.5, and their equations are in Table 2.2.

NSPI exploits the rapid increase in the reflectance spectra of PV modules (see Fig. 2.5) caused by C-Si absorption. C-Si PV modules include poly-C-Si and mono-C-Si. Both types exhibit similarly decreasing energy absorption capabilities from 600 *nm* to 1150 *nm* (Schinke et al., 2015; Deng et al., 2017), resulting in an increase in reflectance. As the market share of silicon solar cells currently exceeds 90% (Silvestre et al., 2018), this feature can be treated as another main spectral feature of PV modules.

The index of average reflectance in the VNIR (aVNIR) (see Fig. 2.5) targets strong spectral absorption features in the VNIR region of PV materials. Common and traditional PV modules have a low average reflectance between 500 nm and 1000 nm (Czirjak, 2017), which is used to absorb more solar energy in this spectral range. For example, monocrystalline PV

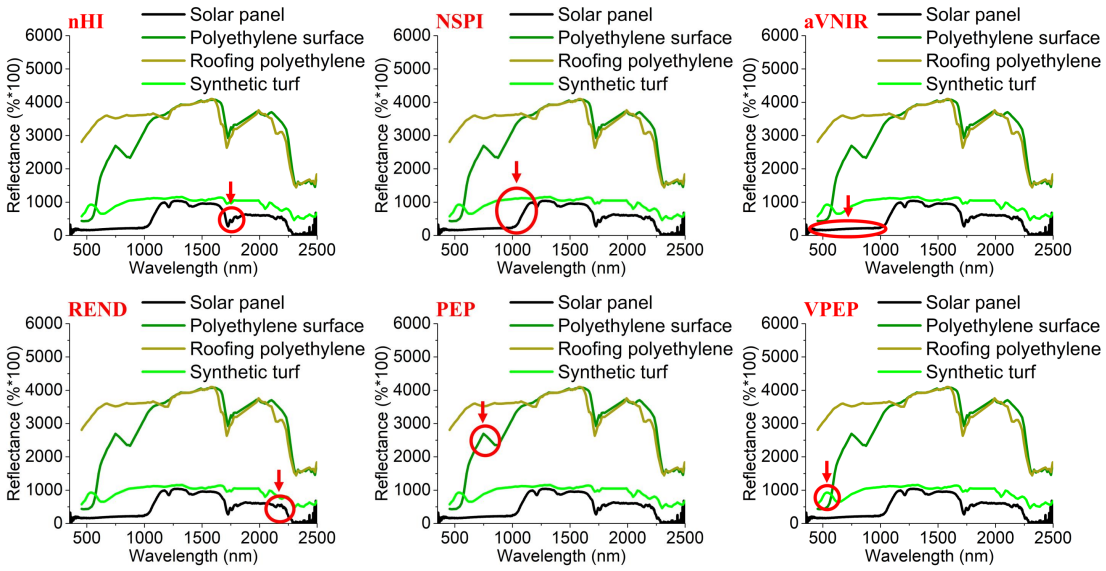


Fig. 2.5 Demonstration of six spectral indices. The PV spectrum is from laboratory spectra-goniometric measurement (monocrystalline PV module, D10). The spectra of polyethylene surface, roofing polyethylene, and synthetic turf are from the HyMap image spectral library.

cells are blackish, and polycrystalline PV cells are dark bluish. In this study, we specified an experimental threshold for aVNIR (see Table 2.2) to primarily eliminate roofing polyethylene.

The index of reflectance drop around 2200 *nm* (REND) (see Fig. 2.5 and see Equation 2.7 in Table 2.2) addresses the typical hydrocarbon absorption properties, since the spectral region from 2200 to 2500 *nm* is affected by numerous overlapping combination and overtone bands. The sheer number of overtone bands causes reflectance to decrease substantially around this region (Herold and Roberts, 2005). These overtone bands can be assigned to the CH_2 and CH_3 stretch and bend, carbonyl-carboxyl C-O stretch, and aromatic carbon stretch (Cloutis, 1989).

The aVNIR and REND (see Fig. 2.5) are physically meaningful spectral indices, but cannot constrain polyethylene surface and synthetic turf misclassified as PV modules. For this reason, the PolyEthylene Peak (PEP) index and the PolyEthylene Peak in Visible range (VPEP) were proposed explicitly for polyethylene surface and synthetic turf, respectively. The PEP feature of polyethylene surface is caused by strong spectral absorption of polyethylene surface at 800 *nm* to 900 *nm* and results in a reflectance peak between 650 *nm* and 860 *nm*. The VPEP feature is due to the fact that synthetic turfs often have visual colors, such as green artificial playgrounds and red artificial running tracks.

Table 2.2 Spectral indices developed and used in the study.

Details	Expression (wavelength/*nm*)	Threshold
normalized Hydrocarbon Index (nHI)	$nHI = \frac{R_{B'} - R_B}{R_{B'}}$ (2.4)	>0.18
Normalized Solar Panel Index (NSPI) (Czirjak, 2017)	$\mathrm{NSPI}_{\mathrm{HySpex}} = \frac{R_{1153} - R_{991}}{R_{1153} + R_{991}}$ (2.5)	>0.15
average reflectance in Visible and Near Infrared Range (aVNIR)	$aVNIR = Mean\left(\sum_{i=500}^{1000} R_i\right)$ (2.6)	<2000
Reflectance drop around 2200 *nm* (REND)	$R_{2100} > R_{2200} > R_{2300}$ (2.7)	- -
A reflectance peak between 0.6 and 0.8 μm, to differentiate from PolyEthylene surface Peak (PEP)	$PEP = R_{750} - R_{650} - \frac{10}{11}\lvert R_{860} - R_{650}\rvert$ (2.8)	<200
A reflectance peak in visible range (VPEP), to differentiate from synthetic turf	$VPEP = R_{630} - R_{470} - \frac{7}{16}\lvert R_{540} - R_{470}\rvert$ (2.9)	<200

2.3.3 PV mapping on Oldenburg

The ten HySpex imaging spectroscopy flight lines were converted from uncompressed band sequential (BSQ) binary files to LZW compressed interleaved geoTIFF, which reduced the data size by a factor of 15 and allowed the data to be read efficiently in chunks for processing. Six spectral indices were sequentially applied to the geoTIFFs, and the pixels that could pass the criteria of all spectral indices were treated as PV module-covered pixels (see Table 2.2). Some classification errors occurred at class boundaries due to spectral mixing within a pixel. These misclassified areas are small compared to the correctly classified areas. Within a class, there are anomalous pixels due to noise in the data. These areas are small compared to the overall pattern. Since one pixel of PV modules is rare at a spatial resolution of 1.2 m $\times$ 1.2 m

of airborne imagery, it was considered as a noisy pixel. Therefore, we applied morphological filtering (clump classes) to remove these noisy pixels and maintain the border pixels.

2.3.4 Validation

Four subsets were selected to evaluate the accuracy of PV detection, as shown in Fig. 2.10. Each subset covered 301.2 m $\times$ 199.2 m (251 $\times$ 166 pixels) in size and was co-registered to the HySpex data. Subset A is dominated by a PV power plant, subset B is the area where campus and institute are located, subset C is a residential area, and subset D covers an industrial area. We manually collected validation data on these four subgroups on airborne 3K photos in combination with field checks. The 3K photos were collected at a similar time as the HySpex data, and have a spatial resolution of 10 cm. Due to uncertainty in geolocation between the 3K photos and the HySpex data sets, a manual shift of up to one pixel (1.2 m) was applied to the PV mapping images to better match the validation data. The Overall Accuracy (OA), Producer's Accuracy (PA), and User Accuracy (UA) were obtained with a pixel-to-pixel comparison using the confusion matrix while treating PV and non-PV as two classes.

2.4 Results

2.4.1 Dealing with the spectral intra-class variability

The HI and nHI values of the five materials at different detection angles were calculated, interpolated, and shown in Fig. 2.6. Bitumen material A, B, and PVC material show a value close to zero, because of the absence of a distinct hydrocarbon absorption feature at 1.73 μm, and therefore result in HI and nHI values close to zero. Two PV materials have higher HI and nHI values from their hydrocarbon features. For HI values, two PV materials show variation among different detection angles. The closer the angle of reflection directly opposite to the incident light (zenith angle 55°, and azimuth angle 171°), the higher the HI values are acquired. However, their nHI values show a greatly minimized variation among the detection angles.

To statistically compare HI and nHI, we show that both HI and nHI can separate PV materials from three other materials of the spectra-goniometric data set, and we can define threshold values of 0.015 and 0.18 for HI and nHI respectively, as depicted by Fig. 2.7. However, nHI values show greater separability between PV materials and other materials. In addition, Fig. 2.7 also shows that HI values of the two PV materials have larger variation among different detection angles than nHI values of PV materials, which are relatively stable.

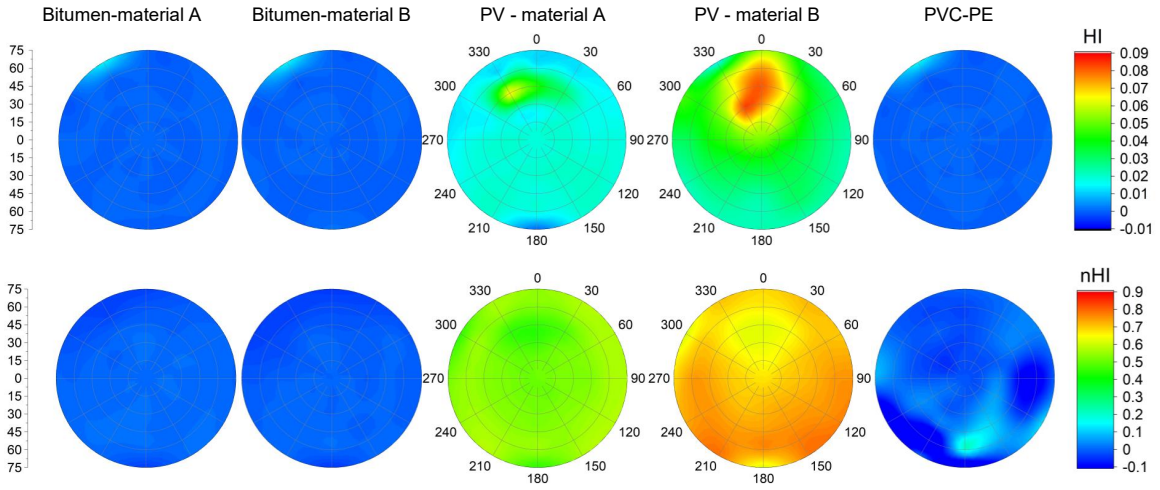


Fig. 2.6 The polar plots of HI and nHI for bitumen material A, bitumen material B, PV material A, PV material B, and PVC with different detection positions. The HI and nHI values were interpolated.

.

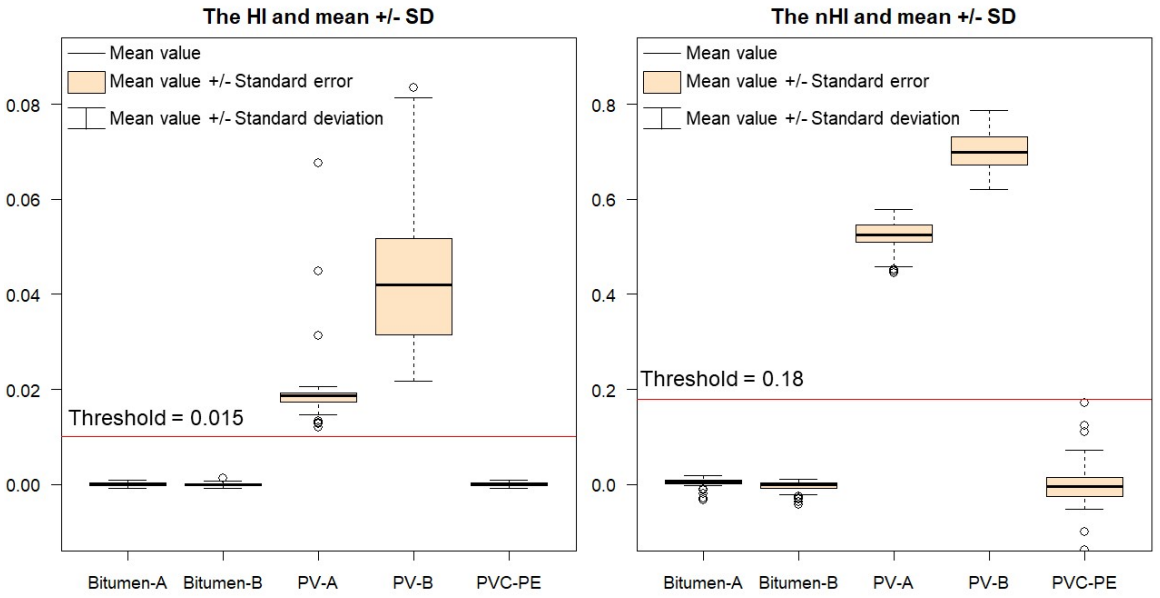


Fig. 2.7 The comparison of HI and nHI with their mean values, ± standard deviations, ± standard errors and outliers.

The HyMap image spectral library was also used to calculate the HI and nHI values (see Fig. 2.8). For the HI values, the spectra of half of the tartan, half of the roofing polyethylene, a few roofing aluminum, roofing concrete as well some pond spectra failed the HI threshold check, meaning that their values were above the thresholds and assumed to be PV. This would

cause difficulties in the following PV detection since these materials could be recognized as PV by HI check. As for nHI, some roofing polyethylene spectra have strong absorption at 1.73 μm, and therefore failed nHI threshold check. In addition, some pond spectra also failed this nHI check. Therefore, although nHI has better performance as it removes the intra-class variability due to different detection angles, it is not good enough to independently distinguish PV from the other surface materials, and some more spectral indices are required to constrain the hydrocarbon-bearing materials.

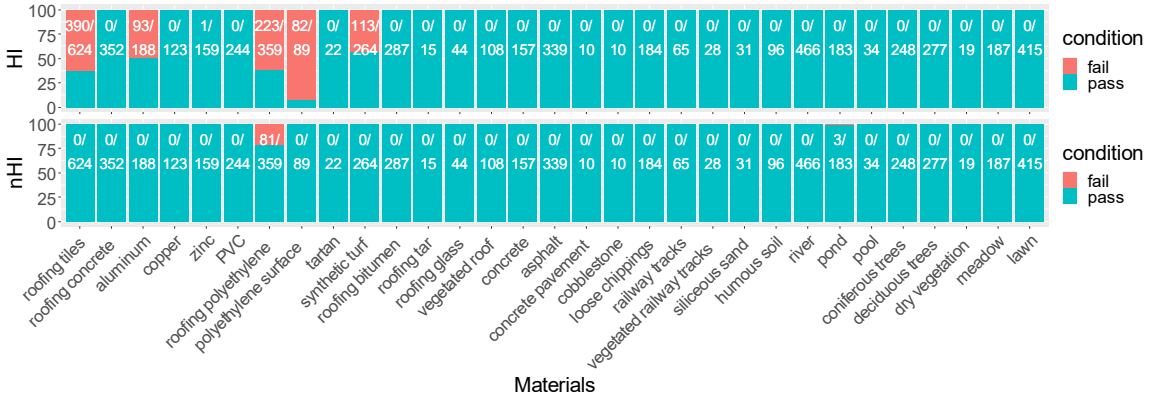


Fig. 2.8 The HI and nHI check of 31 materials in HyMap image spectral library. The ratio between the number of passed spectra and the number of total spectra for each specific material is also displayed.

2.4.2 Dealing with the spectral inter-class similarity

NSPI, nHI, aVNIR, REND were independently applied on the HyMap spectra library, as they can specify different materials based on typical PV features (see Fig. 2.9). NSPI performed quite well, only some spectra of roofing tiles and a few copper spectra failed in this check. The aVNIR had a quite good performance with checking tartan (0 fail), most roofing polyethylene (22/359 fail proportion). REND index independently did not perform well on this check, because the addressed feature is present in most materials. We still keep REND to eliminate false positives caused by image noise. The Sumindices gave sufficient results for this large spectral library and no spectral fail occurred with the combination of these four indices.

2.4.3 PV mapping result

The PV mapping results of the entire study area of Oldenburg were obtained. Fig. 2.10 shows the entire detected PV areas on the left side, about 170,000 pixels or 0.24 km^2. On the

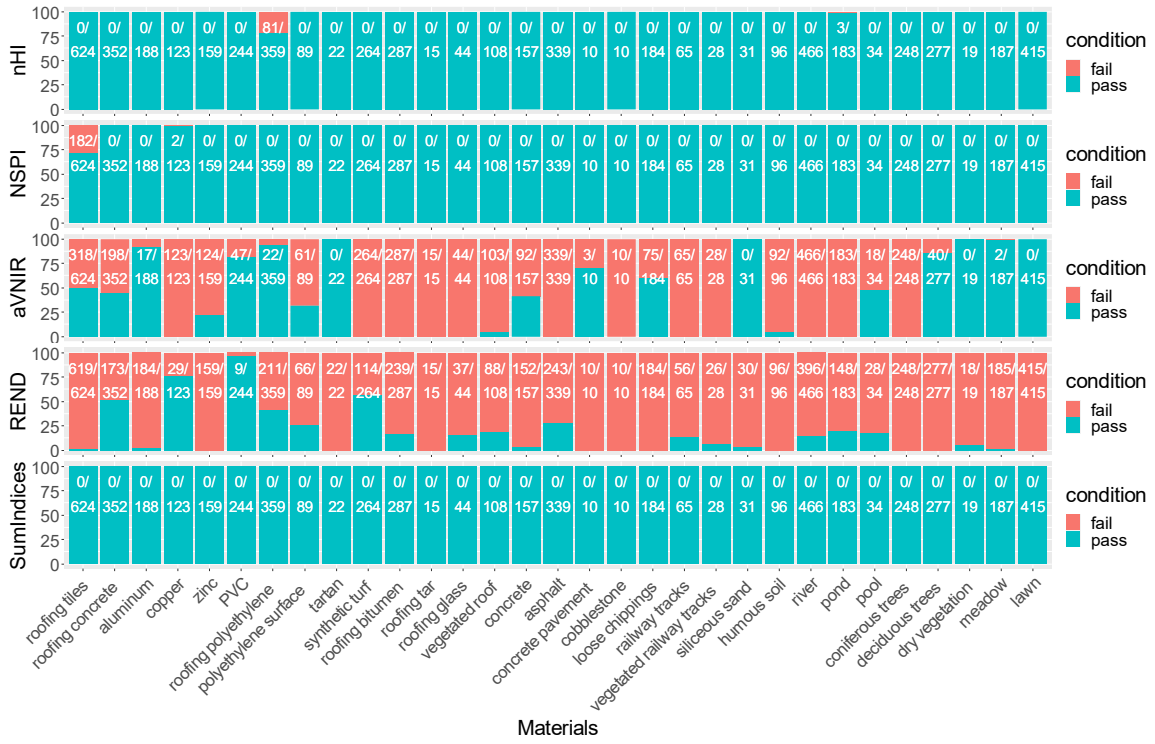


Fig. 2.9 The independent check of four spectral indices (nHI, NSPI, aVNIR, REND) and their combined check (SumIndices) with HyMap image spectral library. The number of failed spectra/number of overall spectra for each specific material is also shown.

right side, four subsets were enlarged. In general, most PV modules were correctly detected from four subsets. Either for the PV power plant of subset A, the campus roofs of subset B, the residential roofs of subset C, the industrial area of subset D, both locations and shapes were correctly detected. Further statistical results were obtained in the following validation process.

2.4.4 Validation

To better compare the reference data and the detected PV areas, the OA, PA, and UA were acquired for four subsets. Together with magnified inlay areas, the validation results were presented in Fig. 2.11.

The Overall Accuracies (OAs) of the four subsets range between 92.8% and 99.3%, indicating that the methodology developed for this study works well. As shown in Fig. 2.11, each PV object was correctly detected except for several panels in subset D. Subset A covers many PV objects, and the detected PV polygons matched well with the reference data under the condition of 1.2 m spatial resolution of HySpex data. The same detection efficiency was

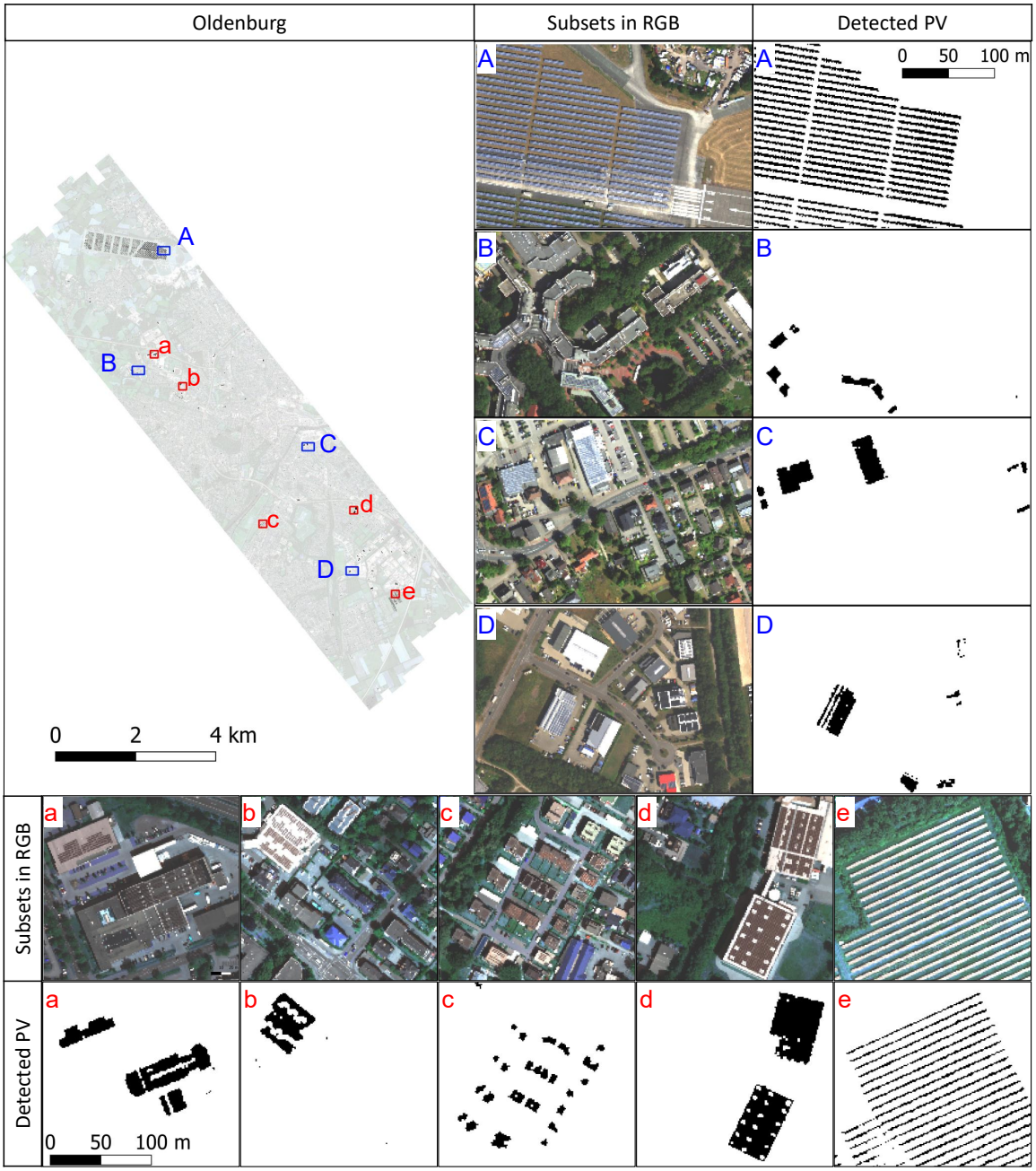


Fig. 2.10 The overall and enlarged detection result in Oldenburg, as the detected PV areas were colored black. Four selected subsets (A, B, C, and D) were marked as blue rectangles in the overall map of Oldenburg, and the RGB and detected PV areas were enlarged on the right. For better illustration, five evenly distributed areas (a, b, c, d, e) were additionally marked in red in the overall map and enlarged at the bottom.

shown in Subset C. Subset B has a PA of 77.8% and a UA of 64.5% since the displacement between the reference data and the HySpex data, as well as the distortion in the reference data

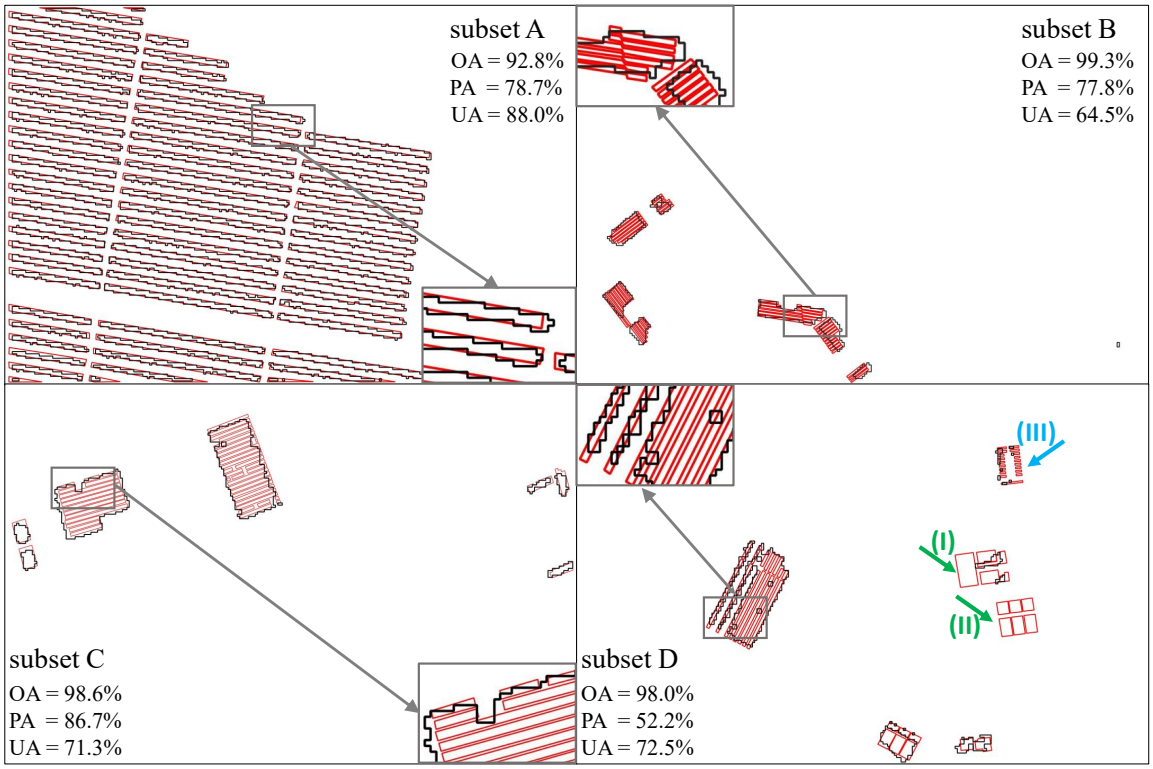


Fig. 2.11 The PV mapping polygons compared with reference data in four subsets. Black polygons show the PV mapping areas, and red polygons show the reference data. The blue and green arrows in subset D show the omission error of the PV detection. To better illustrate the detection accuracy, a region for each subset was selected and depicted within the enlarged inlay figures. For each subset, Overall Accuracy (OA), Producer's Accuracy (PA) and User's Accuracy (UA) were acquired and presented.

from the 3K camera orthorectification. Subset D yielded a PA of 52.2% and a UA of 72.5%, mainly because areas (I) and (II) were omitted from classification as they are thin-film PV modules rather than polycrystalline PV or monocrystalline PV. Furthermore, area (III) was partly missed because the PV modules were too small to be the prevailant spectra at 1.2 m resolution.

2.5 Discussion

2.5.1 PV spectral indices derivation

Our approach exploits the physical absorption and reflectance features of PV modules. Two spectral features present in EVA film and C-Si in PV modules are particularly important for PV detection: The hydrocarbon absorption feature at 1.73 μm is very indicative for

hydrocarbon-bearing materials. Kokaly et al. (2013) discussed the advantage of using the 1.73 μm instead of the 2.3 μm absorption feature in oiled material detection, which holds the same significance in PV detection. First, the 2.3 μm absorption feature can be confused with carbonate absorption in soil, but highly saline soils found in salt marshes are acidic with very low carbonate content. Second, the spectra of dry vegetation also exhibit the 2.3 μm absorption feature (Kokaly and Skidmore, 2015), which are primarily derived from structural biochemical constituents comprising plant cells. Therefore, using the hydrocarbon absorption feature at 1.73 μm as the primary spectral feature to detect PV is a better choice. The second feature is the steep increase in reflectance spectra of PV modules due to the rapid spectral increase of C-Si from 600 *nm* to 1150 *nm* (Schinke et al., 2015; Deng et al., 2017). The NSPI introduced by Czirjak (2017) addresses this feature. As Fig. 2.9 shows, implementing NSPI only is not sufficient to constrain most roofing tiles and a few copper roofs in PV detection. This would lead to confusion since most PV modules were installed on roofs with roofing tiles.

The present approach is able to deal with different detection angles and PV installation angles. The nHI, obtained by normalizing the existing HI, mitigates Bidirectional Reflectance Distribution Function (BRDF) effects from the laboratory spectra-goniometric measurements and shows higher robustness in experimental results. Fig. 2.6 shows the variation of HI values among different detection angles, and nHI removed this variation. The nHI outperforms HI in dealing with material spectral variability and thus offers a better separation. Fig. 2.7 presents the statistical results of HI and nHI values with five materials. Both HI and nHI can easily distinguish two PV materials from two bitumen materials and a PVC sample. However, using HI values to distinguish PV from other materials leads to misclassifications. The nHI has higher separability between two PV materials and other materials. In addition, HI or nHI alone can barely distinguish between monocrystalline PV and polycrystalline PV. Fig. 2.8 shows that nHI has a better performance on the HyMap image spectral library in comparison to HI. HI failed in most hydrocarbon-bearing materials such as most roofing polyethylene, almost all polyethylene surface, and part of synthetic turf, because these materials all have the hydrocarbon absorption feature, and appropriate HI value range is difficult to define. Other than that, HI also failed in checking some spectra of roofing aluminum, roofing tiles, and synthetic turf. nHI only failed in some roofing polyethylene spectra and three pond spectra, leading to more robust results.

A combination of spectral indices is necessary in PV detection because the PV modules are composed of different materials. Using nHI only could cause confusion with hydrocarbon-bearing materials. Still, nHI is a good baseline, and other indices capture additional spectral features of PV or other similar polyethylene materials. Therefore, the combination of HI and

other indices increases the detection accuracy by avoiding false positives. Fig. 2.9 shows that either nHI or the NSPI works well for most surface materials, but not all. Applying nHI alone would misclassify roofing polyethylene as PV modules, and applying NSPI only would misclassify roofing tiles as PV modules.

The approach was trained and assessed with different data types to parameterize the spectral indices, either with respect to different sensors (ASD, HyMap, and HySpex) or different experimental conditions (laboratory spectra-goniometric measurements and airborne imaging spectroscopy data). The large HyMap imaging spectral library was employed to develop other spectral indices to constrain other materials for this study, including 31 surface material classes in 5627 labeled spectra. Therefore, the approach is considered robust to detect PV modules with imaging spectroscopy data from different sensors.

2.5.2 PV mapping results with airborne HySpex imaging spectroscopy data

The approach could quickly and efficiently detect C-Si-based PV modules accurately. In general, Fig. 11 shows the OA of subsets A, B, C and D exceed 90%. If we take a close look at the subsets, only a small section on the east side of subset B was incorrectly classified as a PV module. All other PV installations were correctly detected.

Specially, our approach is able to accurately detect PV modules in different arrangements and within different environments, without the need for explicit training samples for each setting, but purely based on their spectral characteristics. For example, each line of PV modules in the ground-based PV power plant was correctly and accurately detected (see subset A in Fig. 2.11). Had an object-segmentation-based machine learning approach been employed instead, it would have been necessary to sample a sizeable number of labeled training data of PV installations in various arrangements of ground-based settings in addition to samples from residential and industrial rooftops. Hence, a clear advantage of the presented approach is the ability to perform well even in the absence of large sets of labeled training data.

The displacement between hyperspectral data and reference data as well as the distortion in 3K data is the main reason for a UA of 64.5% in subset B, while UAs of other subsets are relatively high. As shown in Fig. 2.11, the PV arrays in the magnified area of subset B were shifted from line arrangement, resulting in the low validation values. This is because the 3K camera is a framing system and therefore affected by relief displacements, which are only partially corrected in the ortho-rectification process. But since a digital elevation model was used, the height of objects above the Earth surface is not taken into account. Therefore, there

are still relative geometric displacements that vary with respect to the building height relative to the Earth surface and additionally increase radially from the perspective center of the 3K image tiles.

Apart from the technical issue, the PA and UA values of four subsets show a limitation of the study. The low spatial resolution of HySpex data of 1.2 m and the resulting over- or under-classification at the edges of PV arrays is the main factor leading to the PA values in the validation. There is a distinct comparison of PA values between subset A and subset C. As can be seen in the enlarged inlay figure (Fig. 2.11), each PV module in subset A and subset C were detected. However, subset C has a PA of 86.7%, while subset A has a PA of 78.7%. This is because subset A contains more objects and therefore more edges were included. The area (III) in subset D also shows the detection limits due to the spatial resolution of 1.2 m of the HySpex data. Some small PV modules were ignored in this detection, since they were too small and scattered distributed, and therefore only covered by spectrally mixed pixels.

Furthermore, the thin-film PV modules should be considered critically in the application of the approach. Although silicon solar cells dominate the market with a share of more than $> 90\%$, and thin-film PV modules have a small share due to low efficiency (Silvestre et al., 2018), it should be noted that thin-film PV modules could not be detected by our approach. In subset D, areas (I) and (II) were not identified (Fig. 2.11) because their spectra are very different from silicon PV modules. Thin-film PV modules are made of different layers and exhibit non-specific characteristics in spectra. Clear identification and differentiation from other materials would require measuring it in the laboratory to explore their spectral reflectance characteristics. Since this PV module is quite rare, it was not available for this study.

In summary, the quality of the PV panel identification is very high (high OA). The lower PA and UA are mainly due to the low spatial resolution of the HySpex data as well as the geometric displacement between the validation and HySpex data.

2.5.3 Future directions

Although the robust approach could be transferred and applied to data collected by different sensors, the band selection of spectral indices would vary slightly. In particular, nHI and NSPI, as the dominant and most efficient indices for PV detection, are quite sensitive to the spectral bands of different sensors they are applied to with their thresholds. Our study developed and combined these spectral indices for the first time, and trained their thresholds using laboratory spectra-goniometric and HyMap spectral libraries, then applied these thresholds with HySpex data on Oldenburg, and acquired accurate PV distributions. Therefore, it is demonstrated that these thresholds are relatively robust for sensors with different spectral resolutions. However,

when the given data is only from a single sensor, adjusting the thresholds of these spectral indices, especially nHI and NSPI, could help to increase the detection accuracy. In particular, nHI and NSPI, as the dominant and most efficient indices for PV detection, are quite sensitive with their thresholds to the spectral bands of the different sensors to which they are applied. Future research could apply machine learning methods to define the thresholds for each spectral index based on training with massive pure spectra.

The physics-based approach presented in this study yields the potential to include new PV materials such as the new thin-film PV modules that could not be detected, because samples were absent in the used spectral libraries. In a first step, these characteristics need to be investigated in a laboratory and tested in airborne or UAV-borne data to then derive robust spectral features that can be parameterized for a broader application. A large amount of training data across different regions is not necessary if the spectral variability can be detected. This makes the physics-based approach a robust and practical method for PV detection.

Detecting large PV modules regionally or nationwide with spaceborne imaging spectroscopy data is efficient and useful in energy system modeling. Currently, the ongoing missions, such as the Italian PRISMA mission (Loizzo et al., 2019) and the upcoming German EnMAP mission (Guanter et al., 2015) are promising data sources for large area PV detection. However, since most spaceborne imaging spectroscopy data have a relatively coarse spatial resolution of 30 m x 30 m, PV mapping in urban areas could be challenging. In this context, a building mask would help improve the mapping accuracy when the detection target is PV systems on roofs. However, results show that the detection of large PV power plants outside urban areas works with high precision. Thus, future investigations could focus on monitoring such sites with spaceborne imaging spectroscopy data. Since effects such as PV soiling due to dust or pollen reduce the efficiency of PVs, this is of high importance to PV power plant operators and owners. Moreover, adding texture, shape, or other visual features from data with the high spatial resolution is a great potential for future studies.

Future studies could explore distinguishing different types of PV based on their unique spectra such as the mono- and polycrystalline silicon, cadmium telluride (CdTe), copper indium selenium (CIS), and copper indium gallium selenide (CIGS) modules (Visa et al., 2016; Burduhos et al., 2018).

2.6 Conclusion

PV modules are not a pure material, but a combination of several materials. Therefore, their detection with imaging spectroscopy data should consider a combination of spectral indices

that are able to separate PV modules and spectrally similar materials. The applicable and robust approach proposed in this study was validated on a large database (spectra-goniometric data, HyMap spectra library, and the HySpex imagery for Oldenburg), and yielded accurate PV areas. Furthermore, BRDF effects due to different detection angles in PV detection were observed and addressed by normalizing the hydrocarbon index. Moreover, the spatial resolution of the imaging spectroscopy data should be sufficient to detect PV material as pure pixels.

This study aims to create greater awareness of the potential importance of imaging spectroscopy data for PV identification. As a physics-based approach, it is robust, transferable, and can provide data on PV coverage on a regional or global scale in a short time. The highlighted analysis of the normalized hydrocarbon index could tackle the detection angle problem in PV installations and data acquisition time, which evidently increases the PV detection accuracy.

It should be noted that the present approach was developed to detect the Si-based PV modules with EVA covers and needs to be further refined and updated to detect other PV modules. Future studies that employ spaceborne imaging spectroscopy data to detect large PV power plants modules should focus on monitoring in order to investigate the potential to detect soiling effects that can decrease the efficiency of such PV modules. The robustness of the developed and tested novel physics-based detection approach for PV power plants paves the way for more refined investigations towards PV type differentiation and the analysis of the efficiency of such modules.

Acknowledgments

General

We would like to thank Prof. Gerd Heilscher and Dr. Benjamin Leutner from HS Ulm for providing various PV modules for spectra-goniometric measurements. We also thank Martin Gutwinski for conducting the collection of the spectra-goniometric data. Moreover, we thank Benjamin Leutner from German Aerospace Center (DLR) for reviewing and modifying the evaluation of our approach in comparison to object-segmentation-based machine learning approaches.

Funding

The study was supported by the e-shape project [European Union's Horizon 2020 research and innovation program, grant no. 820852], the China Scholarship Council [grant no. CSC

201806220088], and the project of “Development of a concept for information retrieval about the building stock in Germany with remote sensing” [German Federal Institute for Research on Building, Urban Affairs and Spatial Development within the Federal Office for Building and Regional Planning, grant no. 10.08.17.7-18.13].

Chapter 3

Sampling robustness in gradient analysis of urban material mixtures

Chapter 3

Sampling robustness in gradient analysis of urban material mixtures

IEEE Transactions on Geoscience and Remote Sensing

Chaonan Ji, Marianne Jilge, Uta Heiden, Marion Stellmes, and Hannes Feilhauer

doi://10.1109/TGRS.2020.3040342

Received March 25, 2020; Revised July 14, 2020 and October 5, 2020; Accepted November 12, 2020.

Available online 7 December 2020

Abstract

Many studies analyzing spaceborne hyperspectral images (HSIs) have so far struggled to deal with a lack of pure pixels due to complex mixtures of urban surface materials. Recently, an alternative concept of gradients in urban surface material composition has been proposed and successfully applied to map cities with spaceborne HSIs without the requirement for a previous determination of pure pixels. The gradient concept treats all pixels as mixed and aims to describe and quantify gradual transitions in the cover fractions of surface materials. This concept presents a promising approach to tackle urban mapping using spaceborne HSIs. However, since gradients are determined in a data-driven way, their transferability within urban areas needs to be investigated. For this purpose, we analyze the robustness of urban surface material gradients and their dependence across six systematic and three simple random sampling schemes. The results show high similarity between nine sampling schemes in the primary gradient feature space (Pspace) and individual gradient feature spaces (Ispaces). In comparing the Pspace to the Ispaces, the Mantel statistics show the resemblance of samples' distribution in the Pspace, and each Ispace is rather strong with high credibility, as the significance level is $P < 0.01$. Therefore, it can be concluded that material gradients defined in the test area are independent of the specific sampling scheme. This study paves the way for subsequent analysis of the stability of urban surface material gradients and the interpretation of material gradients in other urban environments.

3.1 Introduction

The structure and patterns of urban developments have a large effect on urban ecosystem functions (Alberti, 2005). Urban surface conditions strongly influence the climate within cities on the micro scale as well as on the meso- and macro- scale (Oke, 1987). To better understand the climatic interactions within the urban ecosystem, urban surface material compositions are of great interest to climate modelers (Chen et al., 2012). By providing global coverage of pixel-based spectral signatures, spaceborne HSIs can supply substantial information about ecosystem characteristics (van der Linden et al., 2019). However, the

spatial resolution of most operating and future spaceborne HSIs (e.g., EnMAP (Guanter et al., 2015), PRISMA (Lopinto and Ananasso, 2013), DESIS (Alonso et al., 2019; Krutz et al., 2019) and HISUI (Matsunaga et al., 2018)) leads to high spectral mixtures in urban areas (Jilge et al., 2019; Mohanty et al., 2019). Therefore, urban mapping at this spatial scale is still a challenge due to a lack of spectrally pure pixels to train classifiers(Li et al., 2019).

Nevertheless, the mixture within a pixel of urban materials is not arbitrary. In many cases, the actual land use determines the dominant materials that are used for building structures and their surroundings (Niemelä, 1999; Wittig et al., 1998). For example, industrial areas are often characterized by a co-occurrence of concrete, asphalt, and metal roofing; residential areas frequently feature roofing tiles, trees, and lawns. Historical city quarters in Germany such as Wilhelminian style quarters commonly contain roofing tiles, cobblestones pavements for streets, trees, and copper roofs (Heiden et al., 2007, 2012; Heldens, 2010). The distribution of these urban neighborhoods thus leads to reoccurring combinations of material compositions that show strongly inter-correlated distribution patterns (Jilge et al., 2019; Bochow et al., 2007). (Heldens, 2010) The nature of co-occurring surface materials in urban neighborhoods has been the underlying basis for a new technique to analyze urban surface material mixtures. (Jilge et al., 2019) could successfully introduce the gradient concept for urban areas using a test site in Munich, Germany, to relate the distribution patterns of urban surface material compositions to urban neighborhoods. So far, the gradient concept has only been used by vegetation ecologists to characterize floristic gradients for mapping continuous and discrete patterns in plant species assemblages (Schmidtlein et al., 2007; Feilhauer et al., 2011, 2014; Neumann et al., 2016). Treating the mixture of urban surface materials in analogy to the species assemblages in natural vegetation stands, (Jilge et al., 2019) proofed the existence of urban material gradients. The similarity of the surface material composition collected by 153 samples was explored by a feature reduction method (detrended correspondence analysis, DCA). The resulting gradient scores have been regressed against simulated hyperspectral EnMAP data that have a spatial resolution of 30 m x 30 m pixel size. The resulting gradient maps showed pattern of similar urban surface material mixtures that could be linked to specific urban neighborhoods such high density block developments (Wilhelminian style quarter), industrial areas and detached house developments.

However, due to the data-driven property of the gradient analysis, the applicability of this approach for other urban areas needs to be further analyzed. Important aspects of the analyses are the impact of the developed training data on the gradient feature space, the regression model stability and the gradient interpretation.

In this paper, we evaluate the impact of the design of the training data sampling on the resulting gradient feature space. Mainly, we aim to answer the following three questions:

(1) Does the slight movement of systematic sampling affect the material gradients? (2) Are material gradients influenced by using random sampling schemes rather than systematic sampling schemes? (3) To better derive surface material compositions for urban areas, what other possible factors could influence the material gradients? Referring to the sampling scheme of the proof-of-concept study (Jilge et al., 2019), we examine five additional systematical sampling schemes with regular offsets and three random sampling schemes and analyze the impact on the gradient feature space.

The paper is organized as follows. Section II provides a background of the study area and the used data sets. Section III describes the methodology of the experiment, whose experimental results follow in Section IV. A comprehensive discussion is presented in Section V. Finally, Section VI concludes this paper.

3.2 Study Area and Data

3.2.1 Study Area

The study area (Fig. 1) is located in the city of Munich, Germany, the same test site as in the study of Jilge et al. (Jilge et al., 2019). It covers 4.12 km^2 from 48.106°N to 48.133°N and from 11.565°E to 11.632°E. The Isar river divides the study area into two parts. The northwestern part is mainly occupied by old buildings from the Wilhelmine era. In the southeast part, a large vegetated cemetery is situated in the center, surrounded by the Munich East Railway Station, industrial areas, and various residential and commercial areas. The large variety of urban surface materials and their specific compositions are typical for large German cities and provide ideal conditions for investigating the robustness of material gradients among different sampling schemes.

3.2.2 Data

A detailed surface material map (Fig. 1) was used to sample the material composition as the ground truth in order to define the material gradients. The surface material map was prepared by Helden et al. (Heldens, 2010) based on airborne HSI with a spatial resolution of 4 m x 4 m, recorded with the HyMap sensor (Cocks et al., 1998) in a flight campaign in June 2007, and pre-processed by (Richter and Schlaepfer, 2011; Mueller et al., 2005; Habermeyer et al., 2008) . The surface material map initially contained 42 material classes. The modified surface material map removed unlabeled pixels, shadow pixels, facade pixels, and materials which were not covered by every sampling scheme. Finally, 27 urban surface material classes were selected analogous to (Jilge et al., 2019). For more details about the pre-processing of

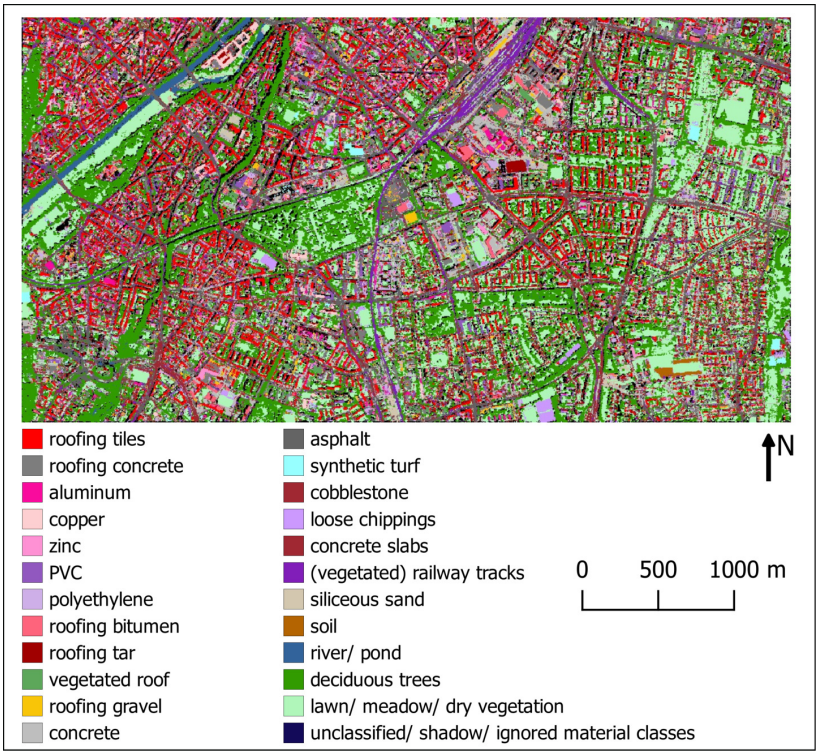


Fig. 3.1 Detailed urban surface material map of the study area in Munich, Germany, determined from HyMap data. Each color represents an urban surface material.

the HSI data and the accuracy of the surface material map, interested readers are referred to (Heldens, 2010; Heiden et al., 2012; Jilge et al., 2019) and the references therein.

3.3 Methodology

A schematic work flow of the proposed approach for sampling robustness analysis of material gradient is shown in Fig. 3.2. In subsection A, six systematic and three simple random sampling schemes are designed. For each sampling scheme, the surface material compositions of each sample are acquired. Subsection B introduces why and how to define material gradients via principle component analysis (PCA). In subsection C, the gradient feature space generated by one sampling scheme ($o4$) is defined as the primary gradient feature space (Pspace). For the visual comparison of the sample distribution in the feature space, all samples from each sampling scheme are projected into the Pspace. The transformation matrix from the original material composition matrix into the gradient score matrix of samples is referred to as loadings. In subsection D, each sampling scheme generate its individual gradient feature space (Ispace). Loadings from Ispaces are compared with each other, and

the multiple linear regressions of material gradients and the corresponding urban surface materials are analyzed for the similarity of Ispaces. In subsection E, the Pspace is further compared with each Ispace by a Mantel test and a Procrustes analysis to check the similarity between the Pspace and Ispaces.

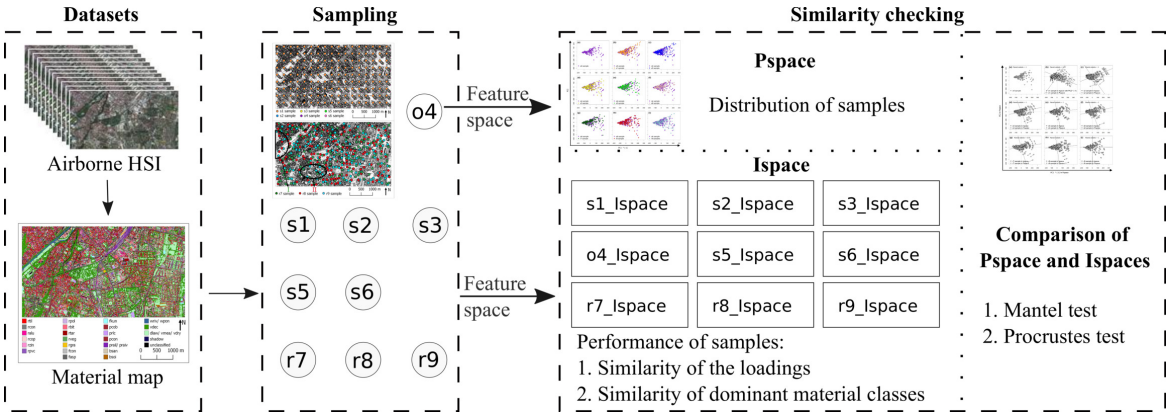


Fig. 3.2 Schematic of the proposed process for sampling robustness of gradient analysis, including (left dashed) data sets, (middle dashed) sampling, and (right dashed) similarity check. $o4$, $s1$, $s2$, $s3$, $s5$, $s6$, $r7$, $r8$, and $r9$ represent 9 sampling schemes.

3.3.1 Sampling design

The sampling robustness of gradients is defined as the transferability and generalizability of material gradients that are applicable to other sampling schemes. For this purpose, this study attempts to assess the similarity of the material gradients generated from different sampling schemes. To analyze the robustness of material gradients, different distributions of samples across the study area are required, which are determined by systematic and simple random sampling schemes. The systematic sampling scheme is intended to support a transition of surface material compositions between neighboring systematic sampling schemes (e.g., $s1$ and $s2$ in Fig. 3.2). The simple random sampling schemes do not follow any regular permutations or center-to-center distance between samples. Therefore, these sampling schemes are believed to more realistically represent the actual situation of the ground truth distribution in many urban environments.

Fig. 3.3 shows the distributions of all sampling schemes, in which the surface material map is displayed as a gray scale map of 27 surface material classes listed according to Table 3.1. Both samples of systematic sampling schemes and simple random sampling schemes have a diameter of 100 m. A total of 153 sampling polygons per sampling scheme are generated.

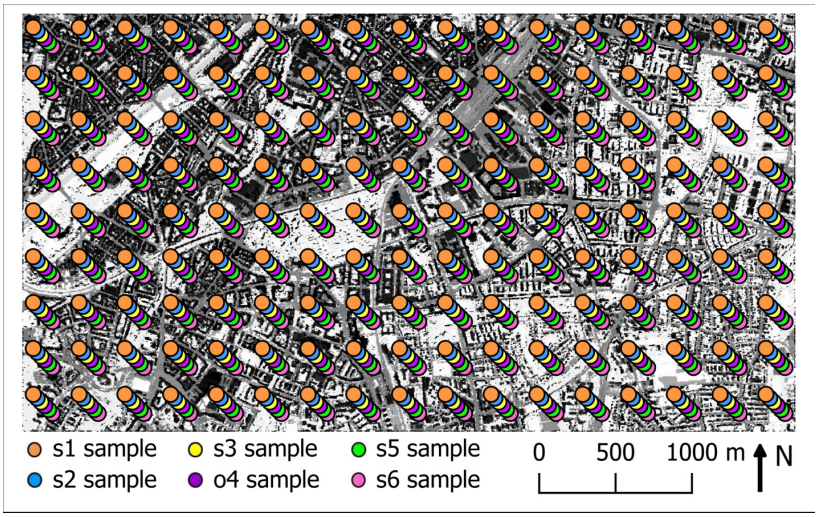
The systematic sampling schemes are evenly distributed over the entire test area, with a center-to-center distance of 300 m between samples (Fig. 3.3a). The systematic sampling scheme used in Jilge et al. (Jilge et al., 2019) is named as $o4$ and marked by purple color. Six systematically chosen sampling schemes are referred to as $s1$, $s2$, $s3$, $o4$, $s5$, and $s6$, respectively. The six sampling schemes are equidistantly spaced by 25 m in both N-S and W-E directions, i.e., about 35.36 m in NW-SE direction (Fig. 3.4), which is close to the spatial resolution (30 m x 30 m) of most spaceborne HSIs (e.g., EnMAP).

Samples of three simple random sampling schemes are randomly distributed in the test area (Fig. 3.3b). Similarly, these three randomly chosen sampling schemes are named $r7$, $r8$, and $r9$, respectively.

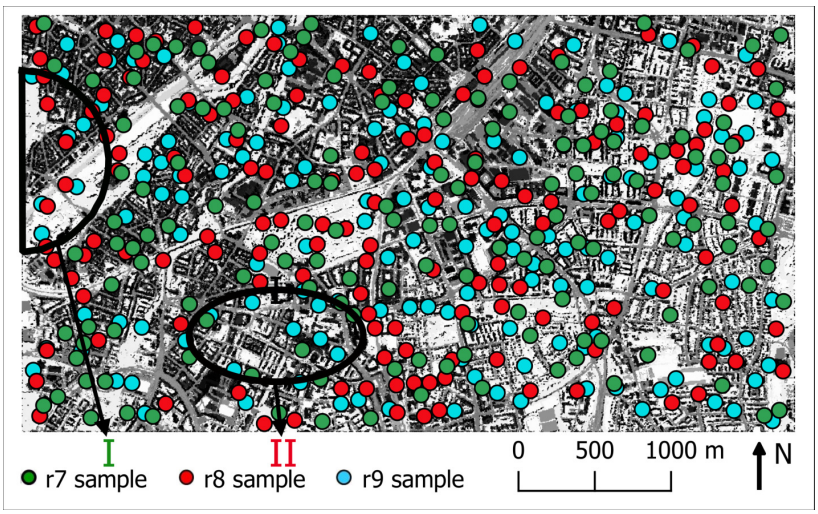
3.3.2 Derivation of material gradients with PCA

The extraction of material gradients is basically a dimensionality reduction of a sample-by-material matrix, and thus, can be performed by many methods (Kruskal, 1964; Green et al., 1988; Tenenbaum et al., 2000). Although the resulting material gradients show generally similar characteristics, the available methods differ in their principles, performance, and the fine details of the extracted material gradients. In Jilge et al. (Jilge et al., 2019), material gradients were defined by DCA from the ground sample in the city of Munich, Germany. DCA is often used for gradient analysis in ecology, where studies have analyzed and mapped gradual transitions in the composition of plant species (Schmidtlein and Sassin, 2004; Feilhauer and Schmidtlein, 2009; Feilhauer et al., 2011). It assumes that materials have a unimodal distribution along the extracted gradients and is therefore well suited for gradients with a pronounced turnover in material composition (Hill and Gauch, 1980). With DCA, however, it is difficult to compare resulting gradients with any other gradient feature space, because new observations can not simply be projected into an existing gradient space (Jilge et al., 2019; Knox, 1989). In addition, the study (Jilge et al., 2019) has shown that the Munich area does not feature a full material turnover and thus, the unimodal distribution model is not required to describe the material gradients in this area.

In this study we chose PCA to extract material gradients, as it is based on a linear distribution model and allows for a better comparability of different material gradients (Hotelling, 1933). PCA is a widely used dimensionality reduction technique (F.R.S, 1901; Jing Wang and Chein-I Chang, 2006; Li et al., 2012b), which reduces the dimensionality of data and at the same time retains most of the variations presented in the data set (Hotelling, 1933; Jolliffe, 2011). It achieves the reduction by transforming the data into a new set of variables called principal components (PCs). The PCs are linear combinations of the original variables and are referred to as gradients in terms of their physical significance. The gradients



(a) Six systematic sampling schemes ($s1$, $s2$, $s3$, $o4$, $s5$, and $s6$)



(b) Three simple random sampling schemes ($r7$, $r8$, and $r9$)

Fig. 3.3 Distributions of systematic and simple random sampling schemes. Each sampling scheme is displayed in one color, and circles represent the size and position of samples. Area I is not covered by the $r7$ sampling scheme, and Area II is not covered by the $r8$ sampling scheme. The enlarged systematic sampling schemes are shown in Fig. 3.4

.

are uncorrelated and hierarchically ordered, so that the first few gradients retain most of the variation presented in all original variables (Plaza et al., 2005). The number of meaningful gradients is determined according to the broken-stick model, which is considered a stable approach to determining the number of PCs (Frontier, 1976; Jackson, 1993).

Subsequently, samples can be projected on the PCs according to their material compositions and the PC loadings to visually assess the similarities and differences among samples.

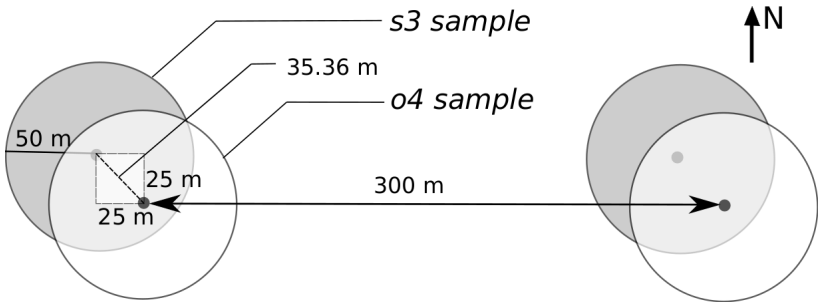


Fig. 3.4 Enlarged systematic sampling schemes. White circles represent samples from *o*4 sampling scheme, and gray circles show samples from *s*3 sampling scheme. The spacing between adjacent schemes and adjacent samples is shown as a dashed line and a double arrow line between circle centers.

Based on this analysis we may determine whether the samples can be grouped (Ringnér, 2008). As the input data here are fractional numbers of material classes, they are fully comparable and for PCA no scaling or centering is needed.

3.3.3 Projection of all samples into the Pspace

To get a first impression of the similarity of all sampling schemes, the samples from each sampling scheme were projected into the Pspace. This projection was made with the loadings generated from *o*4 samples (Ringnér, 2008). The *o*4 samples were selected to generate the Pspace to be consistent with the sampling scheme used by (Jilge et al., 2019), and because they are ordinary and systematically chosen samples. At first, the Pspace is generated by *o*4 samples with PCA, and includes the first and the second gradients (PCs) according to the broken-stick model. The *o*4 loadings are then extracted by the transformation from the original material composition matrix into the gradient score matrix. Finally, all samples from other sampling schemes are converted into the Pspace by applying the *o*4 loadings.

3.3.4 Sample distribution in the Ispaces

The specific Ispace is the regular gradient space created with PCA by the samples for each individual sampling scheme. Similar to the Pspace, the first and the second material gradients are considered to define each Ispace.

The equation of the PCA is given as follows:

$$A_p \times B_p = C_p \tag{3.1}$$

where A_p is the material matrix of pth sampling scheme; B_p is the loadings generated by pth sampling scheme; and C_p is the gradient scores of samples in pth sampling scheme. The equation (3.1) expands as follows:

$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,27} \\ \vdots & \ddots & \vdots \\ a_{153,1} & \cdots & a_{153,27} \end{bmatrix} \times \begin{bmatrix} b_{1,1} & \cdots & b_{1,27} \\ \vdots & \ddots & \vdots \\ b_{27,1} & \cdots & b_{27,27} \end{bmatrix} = \begin{bmatrix} c_{1,1} & \cdots & c_{1,27} \\ \vdots & \ddots & \vdots \\ c_{153,1} & \cdots & c_{153,27} \end{bmatrix} \tag{3.2}$$

where $a_{i,j}$ represents the ith sample containing $a_{i,j}$ pixels of material class j; $b_{m,n}$ is the element in loadings; and $c_{i,n}$ represents the gradient score of the ith sample in the nth gradient.

This study only takes the first two gradients $C[i,1]$ and $C[i,2]$ in formula (3.1) into account, hence only $B[m,1]$ and $B[m,2]$ are considered here.

Two analyses (Pearson correlation and multiple linear regression) are applied to depict the consistency of material gradients and materials among the Ispaces. Pearson correlation analysis is performed by comparing correlated loadings for each sampling scheme, and evaluated using Pearson correlation coefficient. Multiple linear regression analysis is performed in two steps. First, multiple linear regression is applied between material gradient and all material compositions to acquire the weight of each material. Second, the relevance between the material gradient and highly weighted materials (weight > 0.1) is evaluated by adjusted R^2 from further multiple linear regression.

3.3.5 Comparison of Pspace and Ispace

Pspace and Ispaces are compared with a Mantel test and a Procrustes analysis (Legendre and Legendre, 2012; Mardia et al., 1979; Peres-Neto and Jackson, 2001). Both techniques can be used to compare the mutual sample arrangement in the gradient feature space and check whether two neighboring samples in the Pspace are likewise located adjacently in the Ispace. The Mantel test gives an overall estimate of whether the sample distributions in the Pspace and the Ispace match. The Procrustes analysis provides a more detailed assessment and identifies local distortions by quantifying a residual for each data point. In particular, the Mantel test is based on a Pearson correlation between dissimilarity matrices of samples in the Pspace and Ispaces. The dissimilarity matrix of the samples is derived from the Euclidean distances (Legendre and Legendre, 2012). The significance of the correlations is evaluated by means of permuting rows and columns of the first dissimilarity matrix while a total of 999 permutations are used (Dixon, 2003).

Table 3.1 Material table and abbreviations of 27 material classes, the material coverage of the entire surface material map (Fig. 3.1), and the material coverage in samples of each sampling scheme.

Abbreviation	Surface material	Total material coverage in material map [pixel]	Material coverage in one sampling scheme/ Material coverage in surface material map [%]								
			*s*1	*s*2	*s*3	*o*4	*s*5	*s*6	*r*7	*r*8	*r*9
rtil	roofing tiles	66886	8	8	8	9	9	9	8	8	8
rcon	roofing concrete	27440	9	9	8	7	7	8	8	7	10
ralu	roofing aluminum	10466	10	9	8	8	9	10	8	8	9
rcop	roofing copper	13366	7	7	7	8	8	9	8	9	8
rzin	roofing zinc	7607	8	9	8	8	7	5	9	9	8
rpvc	roofing PVC	13434	7	8	7	8	8	7	8	9	8
rpol	roofing polyethylene	8625	10	12	11	9	8	9	7	10	7
rbit	roofing bitumen	14883	10	10	10	8	6	7	9	9	7
rtar	roofing tar	29249	8	8	8	8	8	8	9	8	8
rveg	vegetated roof	18879	8	9	8	8	8	8	9	8	7
rgra	roofing gravel	8206	10	9	9	11	13	13	7	7	8
fcon	concrete	42104	8	8	9	10	10	8	8	8	8
fasp	asphalt	84854	9	8	8	8	8	8	10	8	8
fkun	synthetic turf	3209	10	5	6	9	11	15	5	22	4
pcob	cobblestone	47358	9	9	8	8	9	8	8	7	7
prlc	loose chippings	20546	7	8	10	10	9	9	7	7	7
pcon	concrete slabs	11015	9	9	8	8	8	9	8	7	7
prail	railway tracks	10811	8	9	8	7	7	7	9	10	5
praiv	vegetated railway tracks	11546	10	10	8	8	7	7	7	9	7
bsan	siliceous sand	11765	8	8	8	9	9	9	8	9	11
bsoi	humus soil	2978	4	5	5	6	10	11	5	8	14
wriv	river	4518	10	10	11	10	7	5	8	9	5
wpon	pond	4691	7	8	9	8	9	8	9	9	8
vdec	deciduous trees	172784	9	8	8	8	8	8	9	8	8
vlaw	lawn	16983	7	8	9	8	8	7	8	7	9
vmea	meadow	87525	7	8	8	9	8	8	8	9	9
vdry	dry vegetation	35690	9	8	9	9	8	8	8	8	10
	total	787418									

3.4 Results

3.4.1 Sampling

Table 3.1 lists the coverage of surface materials per specified sampling scheme. As shown in Table 3.1, each sampling scheme covers all occurring material classes. The cover fractions of common material classes (e.g., *rtil*, *rcon*, *fcon*, *pcob*, *rtar*, *fasp*, *vdec*, *vmea*, and *vdry*) are uniform, mainly varying between 7% and 11%. Nevertheless, cover fractions of several material classes vary substantially among sampling schemes. For instance, *r*9 samples contain only 4% while *r*8 samples contain 22% of synthetic turf in the study area.

3.4.2 Distribution of projected Samples in the Pspace

In a first step, all samples from the nine sampling schemes were projected into the Pspace. Fig. 3.5a shows the distribution of *o*4 samples according to their scores on the first and second material gradients. Along the first material gradient axis, most samples are located at the negative end whereas the number of samples decreases towards the positive direction. The sample distribution ranges between -150 and 250 for the first material gradient, where these values correspond to relative values of the gradient scores. On the second material gradient axis, samples are concentrated around zero, and disperse towards the positive end (about 200) and negative end (about -300). Combining the observations for both material gradients, the samples form a triangular field, with most samples concentrated at the lower left corner, i.e. at the coordinate (-150, 0) and the number of samples gradually decreases towards the other two corners.

Based on the distribution of *o*4 samples, the samples of all other sampling schemes were projected into the Pspace as shown in Fig. 3.5. The diagrams (Fig. 3.5b-i) show that all samples are distributed within almost similar triangle field in the Pspace in terms of range and density.

3.4.3 Correlation of the Ispaces

The Pearson correlation coefficient (F.R.S, 1901) was calculated to compare the loadings of all sampling schemes. The coefficients are listed in Table 3.2. In general, the sampling schemes are characterized by high correlations between their loadings. For the loadings on the first material gradient, the Pearson correlation coefficients range between 0.87 (*s*1 and *o*4, *s*1 and *r*9, and *s*2 and *r*9) to 1 (*s*1 and *s*2) and for the loadings on the second material gradient between 0.72 (*s*1 and *o*4, and *s*1 and *r*8) to 0.99 (*o*4 and *r*9), respectively.

Table 3.2 Pearson correlation coefficient of loadings for the first and the second material gradients.

First gradient										Second gradient									
	*s*1	*s*2	*s*3	*o*4	*s*5	*s*6	*r*7	*r*8	*r*9		*s*1	*s*2	*s*3	*o*4	*s*5	*s*6	*r*7	*r*8	*r*9
*s*1		1	0.97	0.87	0.93	0.97	0.95	0.92	0.87	*s*1		-0.87	-0.81	-0.72	-0.77	-0.8	-0.81	-0.72	-0.73
*s*2			0.98	0.88	0.93	0.97	0.94	0.92	0.87	*s*2			0.97	0.85	0.91	0.95	0.93	0.9	0.86
*s*3				0.96	0.98	0.99	0.98	0.98	0.95	*s*3				0.95	0.97	0.98	0.98	0.98	0.95
*o*4					0.99	0.96	0.95	0.99	0.99	*o*4					0.97	0.94	0.97	0.98	0.99
*s*5						0.99	0.98	0.99	0.98	*s*5						0.98	0.97	0.98	0.97
*s*6							0.99	0.98	0.96	*s*6							0.98	0.97	0.95
*r*7								0.98	0.96	*r*7								0.98	0.96
*r*8									0.99	*r*8									0.98
*r*9										*r*9									

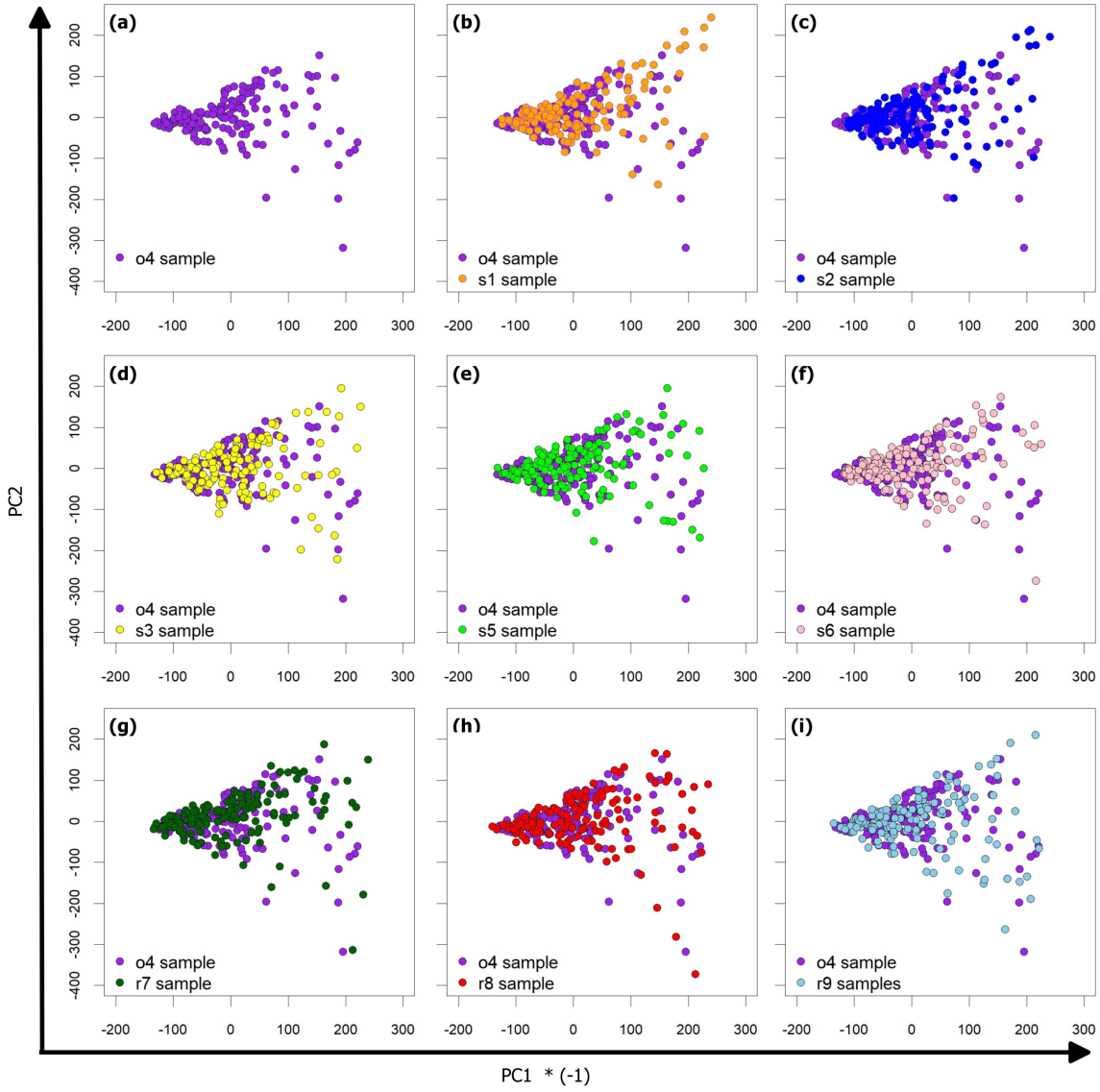


Fig. 3.5 Similar distribution of samples from six sampling schemes in the Pspace. The X axis represents the first material gradient (PC1) in the inverted direction (*(-1)), and the Y axis represents the second material gradient (PC2). The material gradient axes (PC1 and PC2) have no requirements on the direction. PC1 was inverted for a more intuitive interpretation as in (Jilge et al., 2019), where the negative side presents the artificial materials and the positive side shows more vegetated samples. Each sampling scheme is shown in a specific color which corresponds to the distribution of samples in the surface material map (Fig. 3.1).

The regression parameters represent the multivariate weights of the corresponding materials in the derived material gradients in the Ispaces. The higher the absolute regression parameter, the closer the relationship between the material gradient and the corresponding material. The threshold of highly weighted materials is set as 0.1 since it returns the highest adjusted R^2. For instance, *vdec*, *fasp*, *vmea*, *rtil*, and *fcon* are highly relevant with the first material gradient in the *s*1 Ispace (Table 3.3). As expected, all Ispaces are similar in terms of the most relevant material classes to the first material gradient. It can be concluded that the first material gradient from different Ispaces are dominated by these five material classes, though they are produced from different samples. In addition, the second material gradient in *s*1 and also other Ispaces mainly represents *vmea*, *fasp*, *vdec*, *vdry*, and *rtil*, which were the most frequent material class in Table 3.3.

To verify these findings, the regression model was rebuilt between the material gradients and their driving material classes. The resultant adjusted $R^2 = 0.92$ confirms the high relevance between the first material gradient and dominant classes (*vdec*, *fasp*, *vmea*, *rtil*, and *fcon*). Similarly, the second material gradient was regressed with the corresponding five classes (*vmea*, *fasp*, *vdec*, *vdry*, and *rtil*) and an adjusted $R^2 = 0.98$ was returned.

Table 3.3 Multiple linear regressions between the two material gradients and highly weighted urban surface materials for nine sampling schemes.

	*s*1		*s*2		*s*3		*o*4		*s*5		*s*6		*r*7		*r*8		*r*9	
First gradient	*vdec*	-0.93	*vdec*	-0.93	*vdec*	-0.84	*vdec*	-0.68	*vdec*	-0.77	*vdec*	-0.83	*vdec*	-0.79	*vdec*	-0.76	*vdec*	-0.69
	fasp	0.23	*fasp*	0.19	*vmea*	-0.35	*vmea*	-0.57	*vmea*	-0.44	*vmea*	-0.34	*fasp*	0.4	*vmea*	-0.49	*vmea*	-0.58
	vmea	-0.15	*vmea*	-0.17	*fasp*	0.23	*rtil*	0.26	*fasp*	0.28	*fasp*	0.3	*vmea*	-0.37	*fasp*	0.27	*fasp*	0.31
	rtil	0.13	*fcon*	0.15	*rtil*	0.19	*fasp*	0.26	*rtil*	0.26	*rtil*	0.22	*rtil*	0.18	*rtil*	0.21	*rtil*	0.2
	fcon	0.12	*rtil*	0.14	*fcon*	0.18	*fcon*	0.19	*fcon*	0.19	*fcon*	0.15	*fcon*	0.11	*fcon*	0.14	*fcon*	0.13
					rtar	0.1	*rtar*	0.11	*pcob*	0.11	*rtar*	0.1			*rtar*	0.1	*rtar*	0.1
									rtar	0.1								
Second gradient	*vmea*	0.68	*vmea*	-0.88	*vmea*	-0.86	*vmea*	-0.75	*vmea*	-0.78	*vmea*	-0.83	*vmea*	-0.85	*vmea*	-0.81	*vmea*	-0.72
	fasp	-0.62	*vdec*	0.26	*vdec*	0.44	*vdec*	0.64	*vdec*	0.54	*vdec*	0.45	*vdec*	0.48	*vdec*	0.56	*vdec*	0.66
	vdec	-0.26	*fasp*	0.24	*rtil*	0.13			*rtil*	0.23	*rtil*	0.23	*fasp*	0.13	*rtil*	0.13	*vdry*	-0.14
	vdry	0.24	*vdry*	-0.19	*vdry*	-0.12			*vdry*	-0.14	*vdry*	-0.15						
			rtil	0.15	*fasp*	0.11			*fcon*	-0.12	*fasp*	0.12						
			wriv	-0.14														

3.4.4 Comparison of the Pspace and Ispaces

The Procrustes test, the comparison of sample distributions in the Pspace and Ispaces, is shown in Fig. 3.6. Since the Pspace is the Ispace of the *o*4 samples, the distribution of *o*4 samples is congruent in both gradient spaces with a Mantel *r* statistic of 1. Whereas the sample distributions of *s*5, *r*8, and *r*9 sampling schemes only show slight differences, the sample distributions of *s*1 and *s*2 exhibit evident variations (Fig. 3.6) in the transformation from Ispace to Pspace. In general, all transformations from each Ispace to Pspace can be

traced back, and all transformations are not significant. Whereas the sample distributions of *s*5, *r*8, and *r*9 show only minor differences in the transformation from Ispace to Pspace, the sample distributions of *s*1 and *s*2 show a more pronounced dissimilarities (Fig. 3.6). Overall, the samples show a similar distribution in the Ispaces and Pspace, and the transformations do not lead to major distortions. In another perspective, the Mantel *r* statistic of 0.97 between *s*1 Ispace and the Pspace, 0.99 between *r*7 Ispace and the Pspace, and 1 between other Ispaces and the Pspace indicate that there are relatively strong correlations between Ispaces and the Pspace. The p-values of 0.001 acquired by all Mantel test indicate that our results are statistically significant at an alpha of 0.05, representing the high credibility of the results.

3.5 Discussion

The present study aims to investigate whether material gradients are robust among different sampling schemes. In this section, the hypothesis is determined step by step through answering the three research questions.

3.5.1 Are material gradients affected by the slight movement of systematically chosen sampling locations?

The similar distributions of samples from six systematic sampling schemes (Fig. 3.5) provide the impression that the material gradients are robust among slight movement of sample locations in our urban test site.

The high Pearson correlation coefficients (Table 3.2), ranging from 0.87 to 1 between loadings of six systematic sampling schemes for the first material gradient, and from 0.72 to 0.99 for the second material gradient, intuitively demonstrate that the material-to-sample relations from these systematic sampling schemes are similar. Since the loadings enable to project samples from the material composition matrix to the gradient score matrix in the Ispace, the result demonstrates that all Ispaces formed by material gradients generated from the six systematic sampling schemes are consistent. Regarding the multiple linear regressions between the material gradients and urban surface materials (Table 3.3), it is important to point out that the listed materials only indicate the high-weighted materials on the material gradients, which are considered as dominant materials in this study. *vdec* dominates the first material gradient at the first place, while *vmea* dominates the second material gradient for all systematic sampling schemes at the first place. The second and the third dominant materials of the first material gradient are $fasp$ and *vmea* for all sampling schemes except *o*4. Although the second material gradients represent *vdec*, $fasp$, *vdry*, and $rtil$ in different

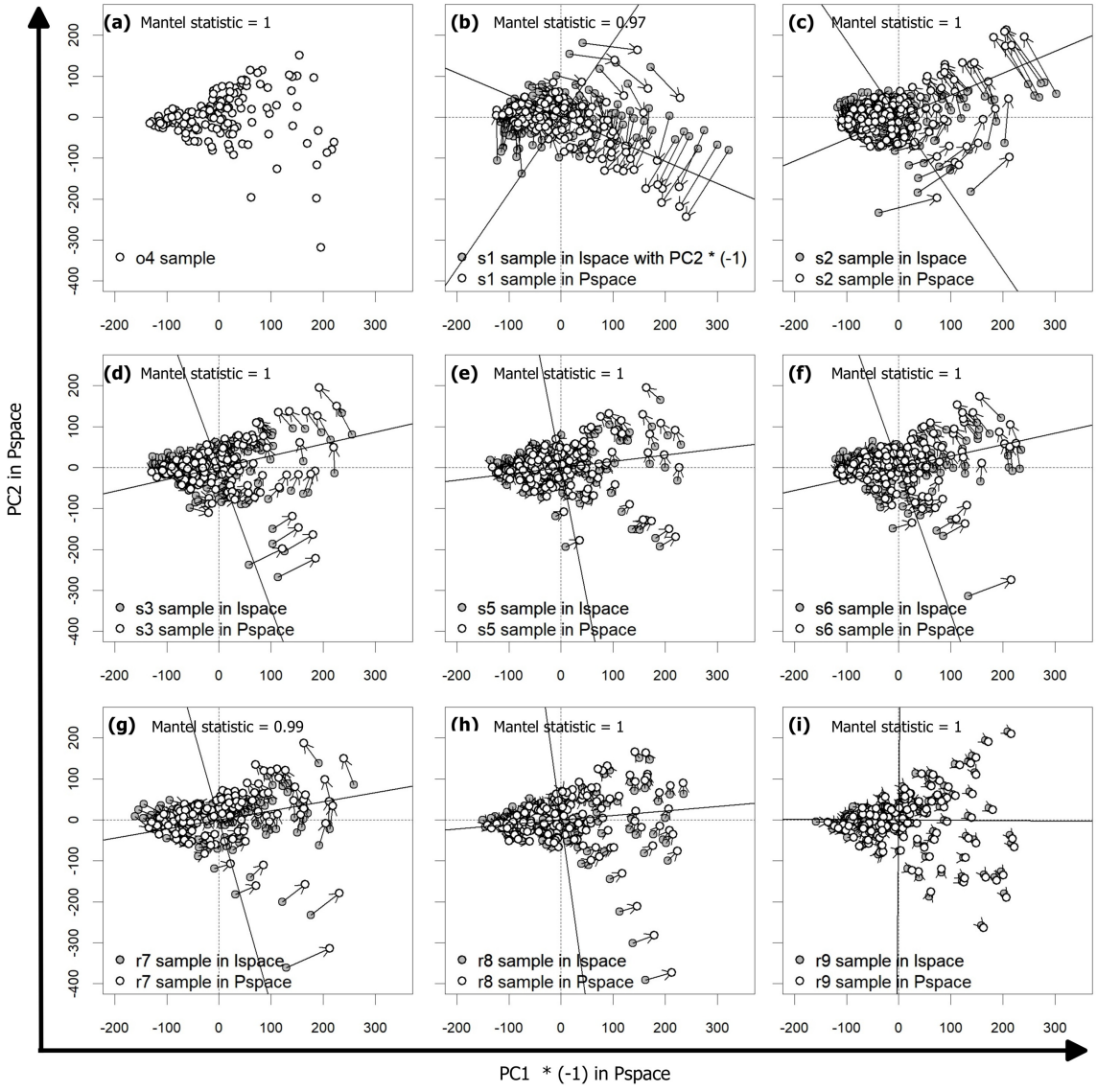


Fig. 3.6 The comparison of the Pspace and Ispaces in Mantel test and Procrustes test. Each sub-figure shows the path of each sample from Ispaces to Pspace (the Procrustes result) and the statistical result (r) of the Mantel test. The Procrustes test finds the configuration change of each sample in each sampling scheme from Ispace to Pspace. Dashed axes represent the gradient axes of Pspace, and solid axes represent the gradient axes of each Ispace. For $s1$ Ispace, the second material gradient is inverted to show the shortest rotation of $s1$ samples from Ispace to Pspace.

weights, the aforementioned results still show that the material gradients generated from different sampling schemes similarly represent 5 materials out of 27 material classes in total.

The results of the Procrustes analysis in Fig. 3.6a-f show that the arrangement of samples in the Pspace is very similar to the distribution of samples in the Ispaces for all six systematic sampling schemes. The results of the Procrustes test show that the Pspace and Ispaces can be matched by rotating the axes. Consequently, the comparison of the Pspace to the Ispaces demonstrates the robustness of material gradients among the systematic sampling schemes.

One reason for the observed gradient robustness is that the material gradients are dominated by the material classes which cover high fractions in the samples, and the changes of materials covering low fractions are ignored. The dominant material classes (*vdec*, *vmea*, $fasp$, $rtil$, $fcon$, and *vdry*) cover both large parts of the study area (Table 3.1) and are distributed evenly. For example, although cobblestone (*pcob*) covers a larger area than concrete ($fcon$) in the study area, concrete dominates the material gradients compared to cobblestone as concrete is distributed more uniformly than cobblestone. Therefore, the slight movement of systematic sampling has no obvious influence on the material gradients.

3.5.2 Are material gradients affected by using random sampling schemes rather than systematic sampling schemes?

The presence of non-stationarity and anisotropy in the spatial data could have a severe impact on the efficiency of systematic sampling (Dunn and Harrison, 1993). Therefore, the simple random sampling scheme was considered to increase the uncertainty of the material cover fraction in the samples, and thus obtain a broader test scope for the robustness of gradient concept.

As discussed in subsection A, all samples including simple randomly chosen samples are similarly distributed in the Pspace. The samples with low values on the negative side of the second material gradient shown in $o4$, $r7$, and $r8$ by visual interpretation are dominated by deciduous trees (*vdec*). Samples with high values on the positive side of the second material gradient shown in $s1$ are dominated by meadow (*vmea*).

The results for first material gradient in Table 3.2 show that the Pearson correlation of loadings between $s1$ and $o4$, $s1$ and $r9$, and $s2$ and $r9$ are lower than others. The results of the second material gradient in Table 3.2 show loadings between $s1$ and other sampling schemes, between $s2$ and $o4$, between $s2$ and $r9$ have Pearson correlation values lower than 0.9. Thus, the findings give no systematic differences between the loadings of systematic and simple random sampling schemes.

Also, as mentioned above, the multiple linear regressions between the two material gradients and urban surface materials are similar for both the systematic sampling schemes and the random sampling schemes. The Mantel statistical r is high for each sampling scheme, for example 0.97 for $s1$, 0.99 for $r7$ and 1 for other sampling schemes. The Procrustes test reveals no differences between systematical sampling schemes and random sampling schemes. It implies that the material gradients are not severely affected by sampling strategies. This can be explained by the variance of each material within the sampling schemes that have been used to generate the material gradients (Table 3.1). The material gradient is thereby as mentioned in subsection A largely determined by materials that cover large parts of the study area and at the same time are distributed homogeneously across the study area. For example, Table 3.1 shows varying cover fractions of synthetic turf ($fkun$) between the sampling schemes from 4% ($r9$) to 11% ($s5$) to 22% ($r8$). However, since $fkun$ covers just a few urban areas (only 3209 pixels), it has negligible influence on the material gradient.

Another reason for the robustness of the material gradients is linked to the fact that the samples seem to represent the entire study area. Thus, small local differences in the sampling locations do not influence the integral generation and the physical significance of the material gradients. For example, as shown in Fig. 3.3, Area I is not covered by $r7$ sampling and Area II is not covered by $r8$ samples. However, their statistical results shown in Table 3.3 show no significant difference.

These observations indicate that Tobler's first law of geography 'everything is related to everything else, but near things are more related than distant things' (Tobler, 1970) also applies to the urban material distribution. This distribution shows consequently a positive spatial autocorrelation that characterizes the clustered distribution of the materials. As long as the sampling schemes represent this clustered distribution sufficiently, the exact location of the samples has no severe influence. Obviously, the spatial scale of this autocorrelation matters. Further testing is needed to analyze these scale dependency of spatial autocorrelation in material distributions across urban areas.

3.5.3 What other possible factors could influence the material gradients?

There are several factors influencing the material gradients that need to be discussed. First, we used PCA to define material gradients because it allows for an easy comparability of material gradients derived by different sampling schemes (see section III). In comparison to non-metric multidimensional scaling (NMDS) (Kruskal, 1964), isometric feature mapping (IsoMap) (Tenenbaum et al., 2000), and other unsupervised data reduction methods, PCA

provides loadings, which can be used to easily and repeatably transform samples from a material composition matrix to a gradient score matrix. Thus, all samples were projected into one gradient feature space (Pspace) for assessing the robustness of gradients determined by different sampling schemes. One drawback of PCA is that it assumes the class-conditional distributions are Gaussian. The distribution in the real observational data are often not Gaussian but strongly multimodal in some cases (Li et al., 2012b). Therefore, PCA might not obtain the best suited material gradients existing in a study area and should only be used for the robustness analysis of material gradients. The PCA-gradient-feature-space generated in this study is not comparable to the DCA-space derived by (Jilge et al., 2019) in scale and number. However, using the same samples, the first material gradient obtained from PCA in this study and DCA in (Jilge et al., 2019) contain similar information on vegetation and artificial material compositions. The second material gradient is different for PCA and DCA. While PCA gradient represents the variety of broad vegetation classes, the gradient obtained by DCA was considered to distinguish urban structures containing different artificial materials.

Second, to map urban material compositions using material gradients, detailed and complete ground truth data are needed that represent the complete variety of material occurrence in the study area. If there is a material not covered by the samples, it cannot be detected or considered in the gradient analysis. If the ground truth data inevitably contains areal errors (also referred to as label noise), the classification accuracy is affected (Jiang et al., 2019). As the material gradient takes a large amount of input data into account, we consider that a significant amount of label noise is acceptable. Nevertheless, the question of how label noise affects the derived material gradient needs to be investigated in future analyses.

Third, the robustness of material gradient may be weaker in specific urban areas due to a more complex or heterogeneous distribution of urban surface materials. If all samples are distributed in a small part of the investigation area, it is difficult to determine the material gradient to represent the whole variety of surface materials in the investigation area. In other words, the samples should sufficiently cover the whole variety of surface materials in the study area (Plourde and Congalton, 2003). Thus, this study suggests that the samples should be evenly distributed over the entire study area.

3.6 Conclusion

In order to obtain the potential strength of material gradient applied in fuzzy pattern description of urban material compositions using spaceborne HSIs, this paper analyzes the robustness of material gradients in an urban area among different sampling schemes, based

on the loadings provided by PCA. The results of this study demonstrate that neither a slight movement of systematic sampling schemes nor simple random sampling schemes largely affect the extracted material gradients. The gradient concept has been proven to be robust for different sampling schemes with the same amount and same size of samples. However, samples shall be well distributed across the study area and need to represent the complete variety of material occurrence in the study area. Future studies should investigate the influence of the size of samples as well as the distance between samples that can also influence the representation of urban surface materials in the gradient space (Jilge et al., 2019). We suggest to use PCA, if different feature spaces need to be compared to test the transferability of the gradient concept. However, if the objective is to best explore the material composition of a specific urban area, other feature reduction methods seems to be more suitable such as DCA or NMDS.

From the methodical point of view, this study proposes a holistic approach to compare the data transferability of dimensionality reduction methods. The shown experiments and measures can be used not only for testing the transferability of the gradient concept, but might also be useful for other data-driven pattern description approaches.

This study provides a first step towards the transferability of the gradient approach to other urban areas despite their data-driven character. This is very important for the further exploration of spaceborne HSI covering urban areas to characterize urban spaces quantitatively without the need for discrete classification and pure material training information.

Acknowledgment

The authors would like to thank the China Scholarship Council for the scholarship (No. CSC 201806220088) that enables the research. We would also like to thank Dr. Wieke Heldens from the German Aerospace Center (DLR) for providing the surface material map of Munich. Furthermore we would like to thank Dr. David Marshall for final proof reading that helped to improve the manuscript.

Chapter 4

Are urban material gradients transferable between areas?

Chapter 4

Are urban material gradients transferable between areas?

International Journal of Applied Earth Observations and Geoinformation 100 (2021) 102332

Chaonan Ji, Uta Heiden, Tobia Lakes, Hannes Feilhauer

https://doi.org/10.1016/j.jag.2021.102332

Received 23 November 2020; Received in revised form 8 March 2021; Accepted 29 March 2021

Available online 25 April 2021

Abstract

Urban areas contain a complex mixture of surface materials resulting in mixed pixels that are challenging to handle with conventional mapping approaches. In particular, for spaceborne hyperspectral images (HSIs) with sufficient spectral resolution to differentiate urban surface materials, the spatial resolution of 30 m (e.g., EnMAP HSIs) makes it difficult to find the spectrally pure pixels required for detailed mapping of urban surface materials. Gradient analysis, which is commonly used in ecology to map natural vegetation consisting of a complex mixture of species, is therefore a promising and practical tool for pattern recognition of urban surface material mixtures. However, the gradients are determined in a data-driven manner, so analysis of their spatial transferability is urgently required. We selected two areas—the Ostbahnhof (Ost) area and the Nymphenburg (Nym) area in Munich, Germany—with simulated EnMAP HSIs and material maps, treating the Ost area as the target area and the Nym area as the well-known area. Three gradient analysis approaches were subsequently proposed for pattern recognition in the Ost area for the cases of (i) sufficient samples collected in the Ost area; (ii) some samples in the Ost area; and (iii) no samples in the Ost area. The Ost samples were used to generate an ordination space in case (i), while the Nym samples were used to create the ordination space to support the pattern recognition of the Ost area in cases (ii) and (iii). The Mantel statistical results show that the sample distributions in the two ordination spaces are similar, with high confidence (the Mantel statistics are 0.995 and 0.990, with a significance of 0.001 in 999 free permutations of the Ost and Nym samples). The results of the partial least square regression models and 10-fold cross-validation show a strong relationship (the calculation-validation R^2 values on the first gradient among the three approaches are 0.898, 0.892; 0.760, 0.743; and 0.860, 0.836, and those on the second gradient are 0.433, 0.351; 0.698, 0.648; and 0.736, 0.646) between the ordination scores of the samples and their reflectance values. The mapping results of the Ost area from three approaches also show similar patterns (e.g., the distribution of vegetation, artificial materials, water, ceremony area) and characteristics of urban structures (the intensity of buildings). Therefore, our findings can help assess the transferability of urban material gradients between similar urban areas.

4.1 Introduction

The majority of the world population lives in urban areas, and the number of urban residents is increasing as more regions are rapidly becoming urbanized (DESA, 2018). Accurate and up-to-date maps are important for modelers to study meteorology (Auer Jr, 1978), climatology (Seto and Shepherd, 2009), and ecology (Lakes and Kim, 2012) and for local authorities to understand the growth dynamics and rapid spatial development of their cities (Cao et al., 2020). However, detailed mapping of urban surfaces is challenging because urban surface materials feature complex spatial patterns, i.e., spatially and spectrally heterogeneous natural and artificial land covers (Chen et al., 2018).

Hyperspectral remote sensing has become an important tool in Earth observation. It extends the number of spectral bands from several or dozens to hundreds, providing a continuous spectrum to identify the materials based on their specific reflectance features (Herold et al., 2004; Heiden et al., 2007). Hyperspectral images (HSIs) that contain a considerable amount of detailed information on land cover and the state of the environment can be used for various applications such as urban modelling (van der Linden et al., 2019), ecological surveys (Degerickx et al., 2018; Skowronek et al., 2018), and geological analyses (Kruse et al., 2003). Spaceborne HSIs can provide global coverage with high temporal resolution to support operational product generation and commercial exploitation of the data, for example, to support economic growth as planned for the Copernicus CHIME mission (Nieke and Rast, 2019) and to support climate-related research, which is one of the goals of NASA's SBG mission (Lee et al., 2015) and the upcoming German EnMAP mission (Guanter et al., 2015). The currently operating missions, such as the Italian PRISMA mission (Loizzo et al., 2019) and the German/US mission DESIS (Alonso et al., 2019), are already delivering data on an operational basis for the development of techniques and scientific data products.

However, the acquisition of spaceborne HSIs with sufficient spectral and spatial resolution, good signal-to-noise ratios (SNRs) and high revisit times is still challenging. Due to sensor design considerations, the rich spectral information in hyperspectral data is often not complemented by extremely fine spatial resolution (Li et al., 2012a). For HSIs with 30 m $\times$ 30 m spatial resolution (e.g., recorded by EnMAP), a large number of surface materials on the measurement scale can be mixed. The resulting mixed pixels reflect the composite spectral response of the contained materials, so the application of per-pixel classifiers to images dominated by mixed pixels may result in inaccurate classification (Plaza et al., 2009).

Gradient analysis appears to be a promising approach for addressing the problem of mixed pixels. Gradient analysis is commonly used in ecology to describe and map natural vegetation by treating all pixels as mixed and to describe and quantify the gradual transitions in the cover fractions of the different species (Schmidtlein and Sassin, 2004; Feilhauer et al., 2011, 2014,

2020; Neumann, 2017). Urban environments contain districts with similar structural and compositional characteristics and thus display the co-occurrences of certain urban surface materials. For example, industrial areas often consist of large low- to medium-rise buildings and predominantly impervious open surfaces, whereas residential areas such as detached housing settlements are likely composed of small low-rise buildings and pervious surfaces such as lawns, meadows and trees (Heldens, 2010). When applying the gradient concept to an urban area, urban material gradients were proposed, and it was then confirmed that such gradients exist in urban space and can be linked to spectral mixtures (Jilge et al., 2019).

However, gradients are generally determined in a data-driven manner. Hence, gradients may be only locally suitable so that additional field data collection will be required, if the gradients are transferred to other unknown areas. Such data collection can be an expensive and time-consuming task. Consequently, an analysis and assessment of the transferability of gradients is crucial for their broader application. As a first step in this direction, Ji et al. (2020) investigated the sampling robustness of gradient analysis with slight movement of the sampling location and different sampling schemes. The influence of such slight movements was marginal, and therefore the next step will be to study the transferability of urban material gradients to unknown areas.

Therefore, the objective of this study is to analyze the area transferability of urban material gradients over two subsets of Munich, Germany. We aim to address the following two questions: (i) Are the urban material gradients transferable between the two study sites? (ii) What affects the transferability of urban material gradients? Our results will provide insights regarding the general feasibility of gradient transfer to urban areas, where either limited or no information on surface material composition is available. As a first step to address this problem, we have chosen two areas in Munich city, which are composed of similar urban neighborhoods with expected similar surface material compositions. Based on the findings, we discussed the potential applications of the transferable urban gradients.

4.2 Study area and data

4.2.1 Study site

The present study was conducted on two subsets of Munich, Germany (Fig. 4.1). The first is located in the south-east of Munich city (48.106°N to 48.133°N, 11.565°E to 11.632°E) and is referred to as the Ostbahnhof (Ost) area in this study because it covers the Munich east train station. The Ost area is considered a perfect urban study area because it consists of complex and typical German urban structures (Heiden et al., 2012), i.e. it represents an inner-city,

densely built-up area with residential and commercial buildings from different epochs. The second study area covers the Nymphenburg Palace and is hence called the Nymphenburg (Nym) area; by contrast, this area is characterized by larger vegetation-covered areas (e.g., Nymphenburg Palace Park, Hirschgarten, and Olympia Park). The Nym area was chosen for the transferability analysis of urban gradients because, on the one hand, it has material classes similar to those of the Ost area, while, on the other hand, it contains a unique set of materials in the palace area.

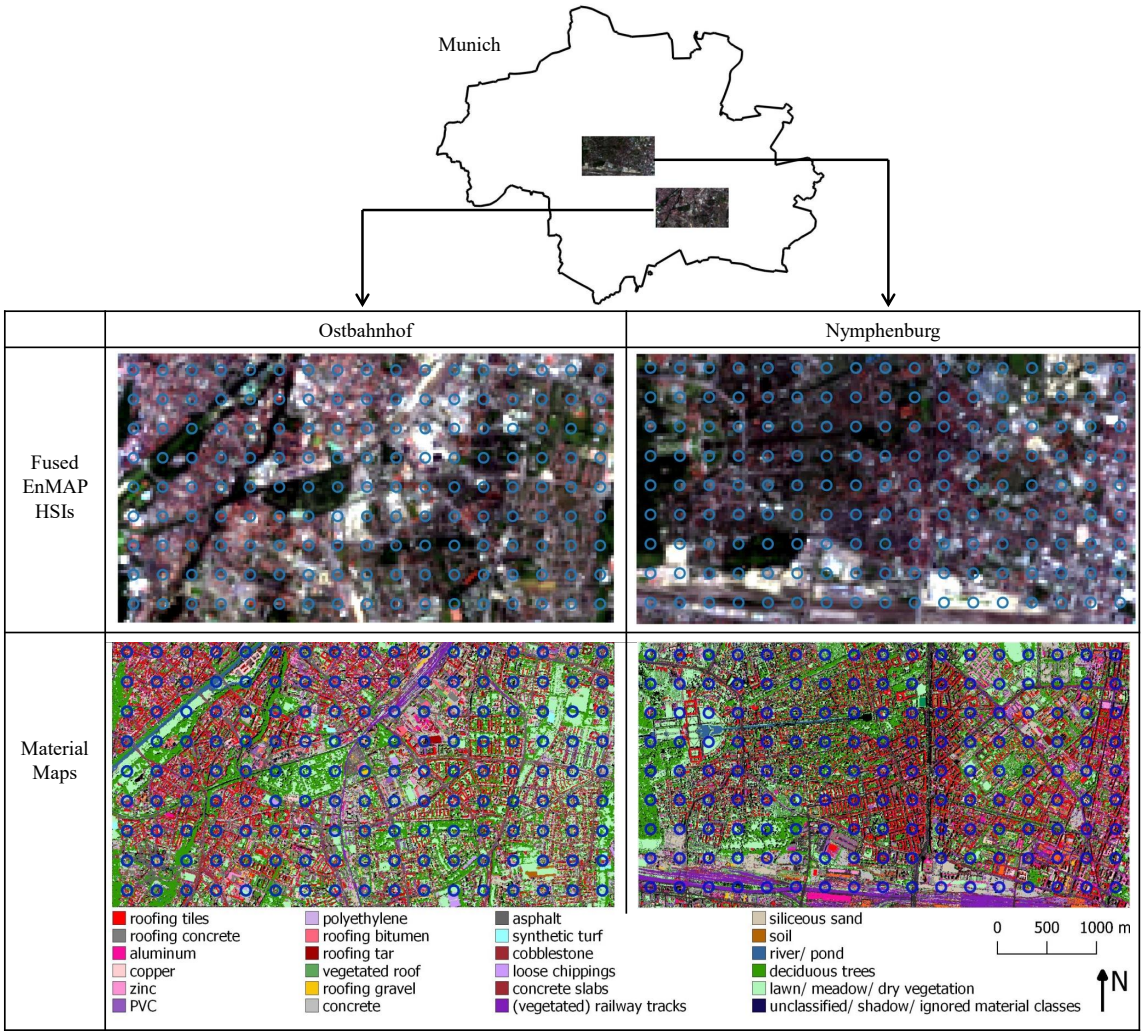


Fig. 4.1 Two study areas: the Ostbahnhof (Ost) area and the Nymphenburg (Nym) area. The simulated HSIs are shown in true colour. The material maps include 27 valid material classes shown in their respective colours and unclassified/shadow/ignored material classes shown in black. The circles show the location and size of the collected samples.

4.2.2 Simulated EnMAP HSIs

The simulated EnMAP HSIs are modelled from the HyMap data acquired by the German Aerospace Center in June 2007 during the HyEurope mission. The HyMap data cover the Ost and Nym areas with two north-south-oriented flight lines. The HyMap imagery includes 128 spectral bands in the range from 450 *nm* to 2500 *nm* with a ground sampling distance (GSD) of 4 m. The data were pre-processed and are reported with an average root mean square error (RMSE) of 0.8 pixels (Heldens, 2010; Heiden et al., 2012). The simulated EnMAP images are characterized by a GSD of 30 m and 242 bands ranging from 423 *nm* to 2439 *nm* (Segl et al., 2012; Guanter et al., 2015). The SWIR data of the EnMAP imagery were taken into account in the overlapping of the VNIR and SWIR sensors. Additional spectral bands ranging from 1358 *nm* to 1418 *nm* and 1814 *nm* to 1951 *nm* were eliminated due to atmospheric water absorption. Therefore, 210 bands of simulated EnMAP HSIs were used in this study.

4.2.3 Material map

The material maps of these two subsets were obtained from a previous study of HyMap HSIs by (Heldens, 2010). The Ost material map was pre-processed by (Jilge et al., 2019) omitting the invalid materials (e.g. shadow, unclassified pixels) that play minor and or no roles in the study area, and thus, 27 material classes were considered. Correspondingly, three additional material classes (lake, pool, coniferous tree) included in the Nym area were omitted because the transferability approach requires the same material classes among different study areas.

4.3 Methods

To test the transferability of the urban gradients, we systematically defined three different conditions, developed the corresponding approaches, and compared the findings for the derived model outcomes and prediction maps. The three conditions were as follows: (i) sufficient Ost samples to produce urban material gradients; (ii) some Ost samples but not enough to perform a gradient analysis; (iii) no Ost samples and therefore no possibility of extraction of the local gradients. We sought to interpret the Ost area under these three conditions, with the assumption that the Nym area provides sufficient samples to generate the urban material gradients. Based on these assumptions, three gradient analysis approaches were constructed as shown in Fig. 4.2. The first approach was used as the control approach (approach-OstOst), and the other approaches are experimental approaches (approach-OstNym and approach-NymNym).

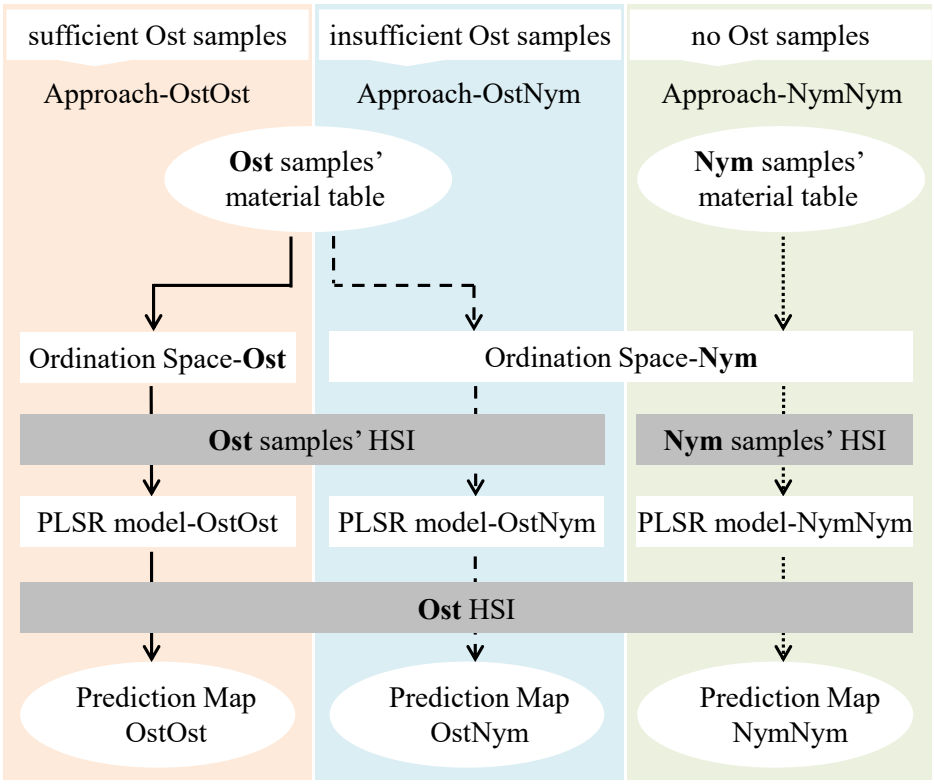


Fig. 4.2 Study workflow: overview of three approaches. Three approaches are proposed to deal with different situations: sufficient, insufficient, and no Ost samples. Ost samples are used in approach-OstOst and approach-OstNym to produce the PLSR models, while Nym samples are used in approach-NymNym. Ost ordination space is used in approach-OstOst, while Nym ordination space is used in approach-OstNym and approach-NymNym.

All analyses were carried out using R Statistical Software 4.4.0 (R Core Team, 2013) and QGIS 3 (QGIS Development Team, 2020). We mainly used the r-packages raster (Hijmans et al., 2013), vegan (Oksanen et al., 2013), autopls (Schmidtlein et al., 2015), and rgdal (Bivand et al., 2015).

4.3.1 Sampling

A total of 153 sampling circles were selected and evenly distributed over each study area, with a diameter of 100 m and a step size of 300 m (Fig. 4.1). The diameter of 100 m was designed to ensure that the sample is large enough to cover material mixtures and to be covered by several spaceborne HSI pixels. The step size of 300 m was used to reduce the effects of spatial autocorrelation in the data (Griffith, 2005; Wang et al., 2012; Jilge et al., 2019). Consequently, the samples fully cover the diverse urban structures dominating the

study site. In addition, each sample was numbered to enable better analysis and discussion later. Table 4.1 presents the material statistics of the samples in the two study areas.

Table 4.1 Statistics of the material map and sampling coverage in the Ost and Nym areas. The abbreviation of materials, total number of pixels for each material class and its proportion in the material map, and sampling coverage pixels for each material class and its proportion in total pixels per class.

Surface Material		**Ost area**				**Nym area**			
	Abbre-viation	**Total Pixels**	**class pixels/ total valid pixels (%)**	**Sampling Coverage**	**pixels in samples/ total valid pixels per class(%)**	**Total Pixels**	**class pixels/ total valid pixels (%)**	**Sampling Coverage**	**pixels in samples/ total valid pixels per class(%)**
roofing tiles	rtil	66886	8.494	6138	9.177	74176	10.371	6929	9.341
roofing concrete	rcon	27440	3.485	2021	7.365	32908	4.601	2802	8.515
aluminum	ralu	10466	1.329	890	8.504	19538	2.732	2002	10.247
copper	rcop	13366	1.697	1149	8.596	6942	0.971	790	11.380
zinc	rzin	7607	0.966	611	8.032	10589	1.481	881	8.320
PVC	rpvc	13434	1.706	1107	8.240	12148	1.699	844	6.948
polyethylene	rpol	8625	1.095	793	9.194	5882	0.822	263	4.471
roofing bitumen	rbit	14883	1.890	1229	8.258	17931	2.507	1186	6.614
roofing tar	rtar	29249	3.715	2585	8.838	7236	1.012	569	7.863
vegetation roof	rveg	18879	2.398	1516	8.030	91709	12.823	7715	8.412
roofing gravel	rgra	8206	1.042	910	11.089	5176	0.724	398	7.689
concrete	fcon	42104	5.347	4275	10.153	37165	5.197	3065	8.247
asphalt	fasp	84854	10.776	6915	8.149	67080	9.379	6082	9.067
synthetic turf	fkun	3209	0.408	291	9.068	675	0.094	32	4.741
cobblestone	pcob	47358	6.014	4046	8.543	37754	5.279	3111	8.240
loose chippings	prlc	20546	2.609	2199	10.703	6733	0.941	588	8.733
concrete slabs	pcon	11015	1.399	954	8.661	1578	0.221	120	7.605
railway tracks	prail	10811	1.373	780	7.215	55175	7.715	4179	7.574
vegetated railway tracks	prailveg	11546	1.466	937	8.115	2280	0.319	123	5.395
siliceous sand	bsan	11765	1.494	1050	8.925	14558	2.036	1550	10.647
humous soil	bsoi	2978	0.378	180	6.044	1528	0.214	95	6.217
river	wriv	4518	0.574	484	10.713	2518	0.352	224	8.896
pond	wpon	4691	0.596	401	8.548	430	0.060	5	1.163
deciduous trees	vdec	172784	21.943	14053	8.133	115810	16.193	9682	8.360
lawn	vlaw	16983	2.157	1471	8.662	37133	5.192	3632	9.781
meadow	vmea	87525	11.115	7774	8.882	43176	6.037	3578	8.287
dry vegetation	vdry	35690	4.533	3165	8.868	7364	1.030	687	9.329
Total Valid Pixels		787418	100			715192	100		
Deleted Pixels		75994				148220			
Total Pixels		863412				863412			

4.3.2 Approach-OstOst

Approach-OstOst applies a gradient analysis following (Jilge et al., 2019) to analyse the Ost area (Fig. 4.2), i.e., the samples collected in the Ost area generated the local urban material gradients using ordination methods and then form the Ost ordination space. Principal component analysis (PCA) was chosen as the ordination method in this study because it achieves the reduction by linear transformation of the data into principal components (PCs, treated as gradients given their physical meaning) and thus allows better comparability of different urban gradients. The first two PCs were selected according to the broken-stick model (Frontier, 1976; Jackson, 1993) in this study. A detailed discussion on the selection

of the PCs in the transferability analysis of gradient and the determination of the number of PCs is found in (Ji et al., 2020). Partial least square regression (PLSR) was used to regress the ordination scores of the samples against their reflectance values retrieved from Ost HSI to train the OstOst PLSR model. Finally, the PLSR model was applied on the Ost HSI to generate the OstOst prediction maps.

Approach-OstOst is a control approach and is expected to produce the most accurate prediction map. In this approach, both samples and ordination space were obtained from the Ost study area, leading to optimized model calibration. This approach serves as a reference in this study and therefore is used to evaluate the other approaches.

4.3.3 Approach-OstNym

Approach-OstNym deals with the situation in which the Ost area provides some samples but the number of these samples is insufficient to perform an urban material gradient analysis, while the samples collected in the Nym area allow a gradient analysis to be performed. In this case, the Ost samples were projected to the Nym ordination space and thus acquired their ordination scores in this ordination space (Fig. 4.2). The OstNym PLSR model was trained by PLS regressing the new ordination scores of the Ost samples against the reflectance values of samples. The prediction maps were obtained by applying the OstNym PLSR model on Ost HSI and are referred to as OstNym prediction maps.

4.3.4 Approach-NymNym

Approach-NymNym deals with the situation in which no Ost samples could be collected, and thus, information from the Nym area was used to interpret the Ost area. In this approach, the ordination scores of the Nym samples in Nym ordination space were PLS-regressed against their reflectance values to generate a NymNym PLSR model (Fig. 4.2). Applying this PLSR model on Ost HSI, the Ost area was interpreted without in situ information.

4.3.5 Comparison of approaches

The three approaches were compared based on the intermediate results obtained from each step. First, the sample distributions in the two ordination spaces were compared visually and statistically using the Mantel test (Peres-Neto and Jackson, 2001) to acquire an overall estimate of whether the sample distributions in the two ordination spaces match. The Mantel test was based on a Pearson correlation in this study between two dissimilarity matrices of the samples' material tables. The Mantel test adopts a permutation test with randomly

permuting rows and columns of the dissimilarity matrix of samples' scores on two PCs 999 times (Legendre and Legendre, 2012) and then recalculates the correlation after each permutation to assess the significance of the observed correlation which is the proportion of permutations that lead to a higher correlation coefficient. In addition to overcoming the problems arising from the statistical dependence of the elements within each of the two matrices, the use of the permutation test means that there is no reliance on assumptions about the statistical distributions of the elements in the matrices.

PLSR models were subsequently generated from the ordination scores of the samples and the sample-averaged reflectance values. It should be noted that the OstNym and NymNym PLSR models were based on the same ordination space—Nym ordination space—with a corresponding set of samples (Ost samples and Nym samples).

Third, the resulting prediction maps obtained from three approaches were visually assessed. When applying the PLSR models to the Ost HSI, three groups of prediction maps were obtained from the three approaches.

In the last step, four groups of samples were selected to demonstrate that a closer location in the ordination space corresponds to more similar material compositions. In addition, the reflectance values were also compared to prove the difference between reflectance values of the Ost area and the Nym area and to prove that the NymNym PLSR model can be applied on the Ost area.

4.4 Results

4.4.1 Ordination spaces

The Ost ordination space and Nym ordination space are shown in Fig. 4.3. Fig. 4.3a presents the ordination space generated from Ost samples that is used for approach-OstOst. Fig. 4.3b shows the ordination space generated from Nym samples that is used in approach-OstNym and approach-NymNym with the assumption that insufficient or no Ost samples generate the ordination space. All of the samples are projected into both ordination spaces, with material vectors representing the directions of increase in the respective material cover fractions.

The distributions of the samples between Ost ordination space and Nym ordination space are similar. Generally, both sample cloud distributions have a triangular shape. A detailed examination shows that most representative material classes are similar. For example, most vegetation types, such as deciduous trees(*vdec*), meadow(*vmea*), and lawn(*vlaw*) are on one side and other materials (mostly artificial materials) on the other. The vector length of vegetation types are longer because they cover more pixels. Furthermore, the distances

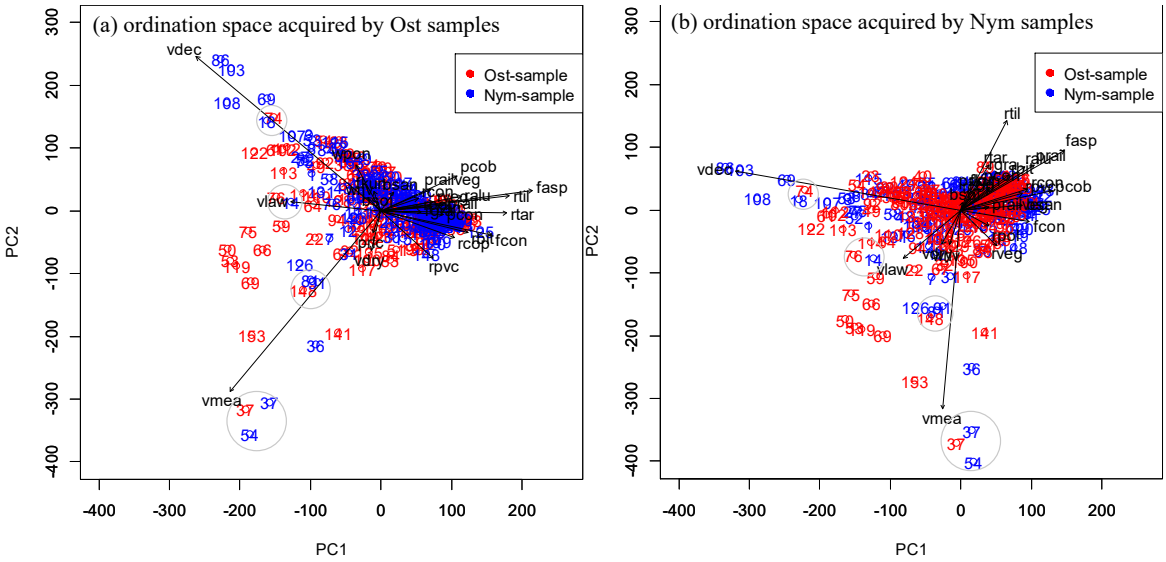


Fig. 4.3 Ordination spaces generated from the samples of the Ost and Nym areas. (a) is used in approach-OstOst, and (b) is used in approach-OstNym and approach-NymNym. Ost and Nym samples are presented, and the potential material vector and their length are also given. The full names of the materials are given in Table 4.1. In addition, the green circles highlight four groups of samples with closely spaced positions.

between the samples are relatively constant. For example, the distances are consistent within the group of Nym-sample-86, Nym-sample-103 and Nym-sample-108, as well as for the group of Ost-sample-37, Nym-sample-37, and Nym-sample-54.

Mantel statistics also show the consistent configuration of the distribution of the samples in two ordination spaces. The Mantel statistical result of the Ost samples is 0.995 with a significance of 0.001 in 999 free permutations, while that of the Nym samples is 0.990, with a significance of 0.001 in 999 free permutations.

4.4.2 PLSR models

We generated six PLSR models by regressing the ordination scores in each gradient and the samples' averaged reflectance values. The best PLSR model for PC1 resulted in R^2= 0.898 for calibration and R^2= 0.892 in 10-fold cross-validation (Fig. 4.4a). Accordingly, the PLSR models of PC1 from approach-OstNym and approach-NymNym also acquire relatively high R^2 (0.760, 0.743; 0.860; 0.836) (Fig. 4.4b, 4.4c). For the PC2, the PLSR models resulted in R^2= 0.433 for calibration and R^2= 0.351 (Fig. 4.4d) in 10-fold cross-validation for approach-OstOst, and for approach-OstNym and approach-NymNym are 0.698, 0.648, 0.736, and 0.646 (Fig. 4.4e, 4.4f).

The distribution of samples in Fig. 4.4 corresponds to the distribution of samples in Fig. 4.3. The approach-OstOst evaluates the ordination scores in the Ost ordination space, and therefore, the distribution of the samples in PC1 is dense and ranges from -200 to 200 in PC2. The approach-OstNym and approach-NymNym apply the ordination scores of the Nym ordination space, and therefore, the distribution in PC1 is loose and ranges from -200 to 100 in PC2.

For approach-OstOst, PC1 (Fig. 4.4a) contains the largest variance, and PC2 (Fig. 4.4d) represents less information and consequently shows a relation to the reflectance values modelled with relatively low R^2. The approach-OstNym and approach-NymNym acquired reasonable PLSR models and performed well with 10-fold cross validation, indicating that the reflectance values of the Ost samples and their ordination scores in the Nym ordination space can reasonably build a PLSR model. In particular, the higher values of calibration and validation of PC2 in approach-OstNym (Fig. 4.4e) indicate that the PC2 in approach-OstNym provides more accurate information than the PC2 in approach-OstOst. Similarly, the relatively high value of calculation R^2 and validation R^2 of the PLSR models of two PCs in approach-NymNym prove that the ordination scores of the Nym samples in Nym ordination space and their reflectance values established an accurate PLSR model.

4.4.3 Prediction maps

Prediction maps were acquired by applying the corresponding PLSR models to the Ost HSI. The prediction map obtained from the PC1 PLSR model of approach-OstOst (Fig. 4.5a) presents the pattern of vegetation in blue and artificial materials in red, and the PC2 prediction map (Fig. 4.5d) presents the different vegetation species with rather low accuracy. The prediction maps generated from approach-OstNym provide similar information: the resulting PC1 prediction map (Fig. 4.5b) vividly displays vegetation coverage, and the structures of artificial materials are similarly indicated; the PC2 prediction map (Fig. 4.5e) also shows the vegetation information. While the prediction maps produced from approach-NymNym appear to lose some detailed information, the major features are provided: the PC1 prediction map (Fig. 4.5c) presents the information of vegetation and artificial materials, while the PC2 prediction map (Fig. 4.5f) shows the vegetation coverage of the study area.

Two interesting phenomena are observed in the prediction maps. The trapezoidal area (A) in Fig. 4.5a, called the Ostfriedhof area, shows variable intensity of the blue colour. The left side is light blue, and the right side is dark blue. In the historical image review by Google at the time closest to the acquisition time of the HyMap data, the area was divided into two sub-areas, with the left side including more graves, i.e., more impervious surfaces, while the right side containing less graves. The trapezoidal area (B) (Fig. 4.5b) shows two types of

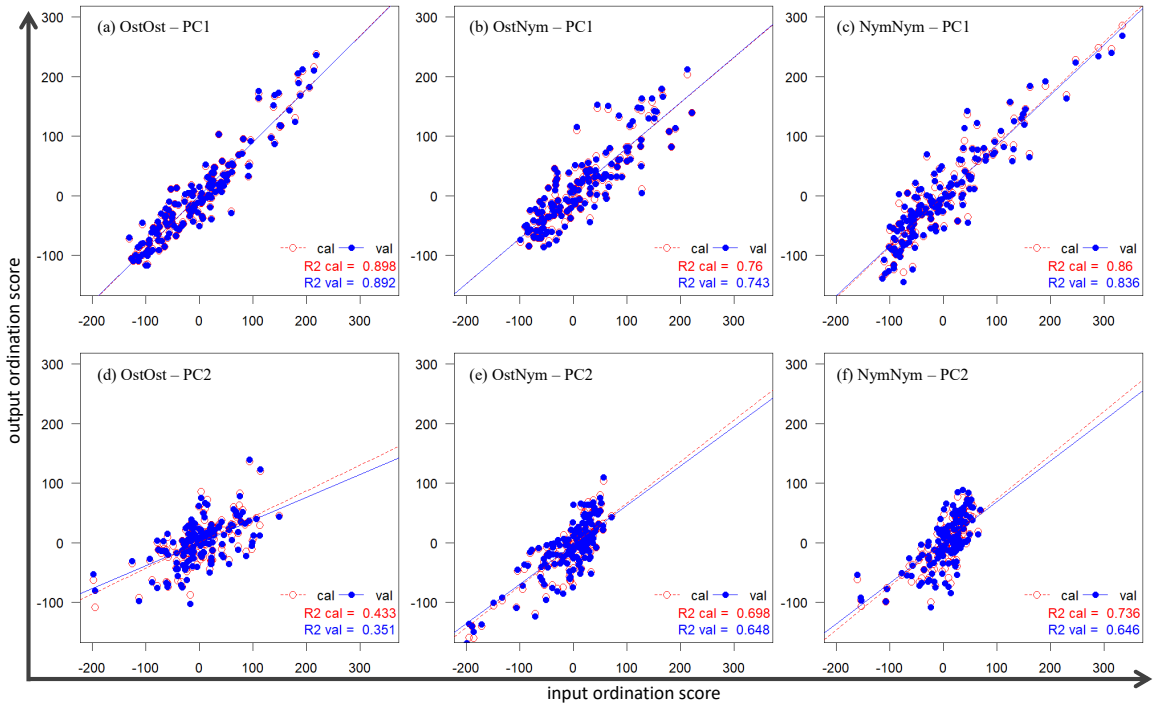


Fig. 4.4 The performance of PLSR models in three approaches. R^2 cal: R^2 in calibration, R^2 val: R^2 in 10-fold validation. Each approach has two PLSR models of PC1-reflectance and PC2-reflectance. Circles or points represent samples. The x-axis represents the input ordination scores of the samples used to build the PLSR model, and the y-axis represents the output ordination scores, while the calculated scores of the PLSR model are represented by red circles and the predicted scores of the 10-fold cross-validation of the PLSR model are represented by blue points. Their fit lines and R^2 are provided and displayed in corresponding colours. $y = x$ represented the best possible fit with either calculated or validated ordination scores of the samples matching the input ordination scores and therefore the best possible calculation or validation R^2 is 1.

patterns with red on the left and blue on the right. Google Earth historical imagery check reveals that the left block contained row houses and the right block contained semi-detached houses. The semi-detached houses had much more vegetation than the row houses and are therefore coloured blue, while the the row houses block is coloured red in the prediction map.

4.4.4 Comparison of material compositions and reflectance values of sample groups

Four groups of samples were selected in the ordination spaces within green circles (Fig. 4.3) for which the reflectance values and material compositions are provided in Fig. 4.6. The first group includes Ost-sample-74 and Nym-sample-18 that have similar material compositions.

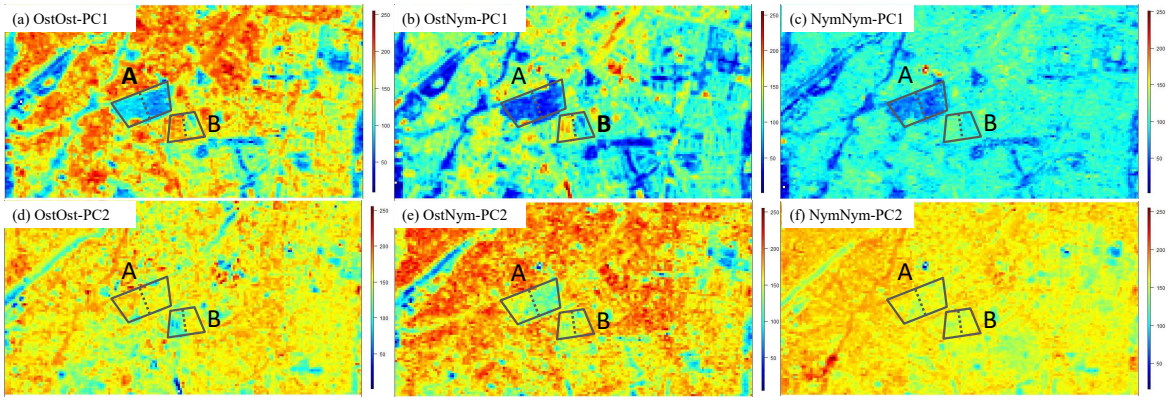


Fig. 4.5 Prediction maps for each PC. Subfigures (a) - (f) show the prediction maps for each PC obtained directly from the PLSR models. For better visual interpretation, PC1 is inverted so that the vegetation pattern is shown in blue and artificial materials in red. On the one hand, three pairs of prediction maps show similar patterns including vegetation coverage and urban structures. On the other hand, detailed information is gradually lost from the prediction maps of approach-OstOst, approach-OstNym, and approach-NymNym.

Since lawn (vlaw) and meadow (vmea) have similar reflectance values, the difference between these two samples is that Ost-sample-74 covers a small amount of asphalt (fasp). Given that the vegetation species have stronger spectral features, their reflectance should not vary much. However, their reflectance values do vary considerably (Fig. 4.6a). The Ost-sample-76 and Nym-sample-14 in the second group are distantly close in the ordination space, and have similar material and same reflectance values, which fits very well with the application of transferable urban material gradients. The other two groups comprise three samples, and show the difference between the reflectance values of Ost and Nym HSIs. Fig. 4.6c displays the material portions and reflectance of Ost-sample-148, Nym-sample-81, and Nym-sample-91. Nym-sample-81 and Nym-sample-91 are closer in the ordination space (Fig. 4.3), and have similar materials (meadow, deciduous trees, roofing vegetation, and cobblestone) with similar proportions, and this is reflected in their spectra. The last group includes Ost-sample-37, Nym-sample-37, and Nym-sample-54 (Fig. 4.6d). The Ost-sample-37 is located in the middle of two Nym samples, but the reflectance values of Ost-sample-37 are the lowest, and the other spectra of the Nym samples are characterized to be more similar. While Fig. 4.6b and 4.6c demonstrate the similarity of the reflectance values of the Ost and Nym areas, Fig. 4.6a and 4.6d show the differences between the reflectance values of these two areas.

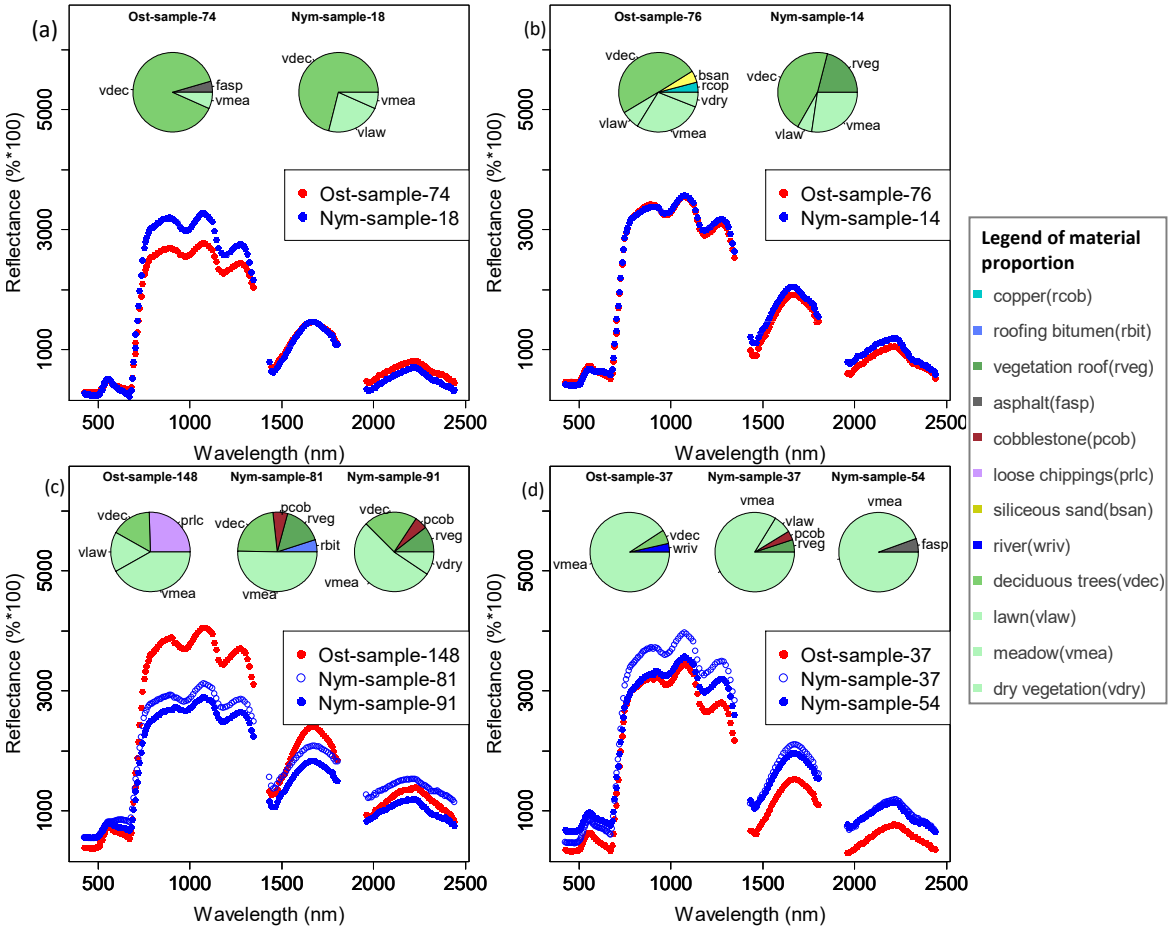


Fig. 4.6 Comparison of reflectance values. Four sets of samples are selected from the ordination space (refer to Fig. 4.3 for detailed discussion. The material compositions of the samples are provided. The material covering less than 15 pixels are neglected in this figure that represents approximately 3% of the total covering pixels of a sample. The color of the material is consistent with Fig. 4.1.

4.5 Discussion

4.5.1 Are the urban material gradients transferable between two study sites?

Theoretically, the transferability of urban material gradients means that the gradients acquired from one area are suitable for the interpretation of another area, i.e., the approach-OstNym, using Nym gradients to interpret Ost area, performs as well as approach-OstOst. The difference between these two approaches is the gradients, i.e. approach-OstOst uses the Ost gradients and approach-OstNym uses the Nym gradients. Other factors, such as the material

composition of the samples and their reflectance values, are fully controlled because the samples used are always Ost samples. As shown in section 4.4, for the sample distribution in the ordination spaces, the performance of PLSR models, and the prediction maps, the approach-OstOst and approach-OstNym exhibit very similar results. Therefore, the urban material gradients are transferable between the Ost and Nym areas.

This means that the material gradients acquired from the Nym area can be used to successfully interpret the Ost area. This is investigated by approach-NymNym and demonstrated the quality of the results through the comparison of approach-OstOst and approach-NymNym. The intermediate results indicate that their ordination spaces are similar and PLSR models performs quite well. However, although the prediction maps of the approach-NymNym retain the main characteristics of the Ost area, they still loses some detailed information.

The comparison of approach-OstOst and approach-NymNym shows that the good performance of the PLSR models does not always mean that both approaches can achieve good interpretation results on the Ost area. The OstOst PLSR model is based on the ordination scores of the Ost samples and reflectance values of the Ost area, whereas the NymNym PLSR model describes the relation between the ordination scores of the Nym samples and the Nym reflectance values. In both ordination spaces (Fig. 4.3), the samples in close proximity to each other have similar material compositions, and thus we can assume that either the Ost sample or the Nym sample can be treated as equivalent. In another words, the samples in close proximity in the ordination space should have qualitatively similar reflectance values. If this is in fact the case, approach-NymNym should display similar information to that obtained by approach-OstOst, i.e. the urban material gradients are transferable under application from the Nym area to the Ost area.

To determine whether the reflectance values vary between Ost and Nym HSIs, four groups of samples were selected in the ordination spaces within the green circles (Fig. 4.3) for which the reflectance values and material compositions are provided in Fig. 4.6. The second and third groups of the samples show the consistent reflectance values between the Ost and Nym samples, while the first and last groups demonstrate that some samples from the Ost and Nym areas do not show consistent reflectance values. Therefore, the PLSR models acquired in approach-NymNym cannot be simply applied to interpret the Ost area, as they are calculated for Nym HSIs, and there is a difference between the reflectance of Ost and Nym HSIs. This phenomenon is discussed further in the next section.

4.5.2 What affects the transferability of urban material gradients?

The gradient concept is based on the assumption that similar material compositions in the gradient space result in similar spectral reflectance mixture characteristics. Therefore, the

consistency of reflectance values from HSIs between the areas are relevant as mentioned above. The neighboring samples in the ordination space (Fig. 4.3a) have similar material composition, and therefore, the difference between the resulting OstNym and NymNym prediction maps shown in Fig. 4.3a is related to the difference of the reflectance values between the Ost and Nym areas. A possible reason for this observation is the underlying data source for the simulated EnMAP data as described in Section 4.2.2. Several airborne HyMap flight lines are combined to generate the EnMAP simulated data set. These flight lines show that differences in the brightness level that still remain can be traced back to the remaining BRDF effects that are then also present in the simulated EnMAP data. It can be expected that real spaceborne HSI data will not show these local brightness differences. However, for this study, real spaceborne HSI data that matches the used surface material map of Munich were not available. Future studies with real spaceborne HSI information such as that from PRISMA or DESIS data are expected to obtain prediction maps with higher accuracy.

Another aspect that affect the transferability of urban material gradients is the material composition of the two areas. Since the PLSR model is trained with a specific set of materials of the first area, new materials in the second area cannot be considered in the PSLR model and therefore, will most likely result in lower model performance. In other words, the new material in the unknown area will not be recognized due to the lack of appropriate input in the training of the PLSR model. Therefore, the detected materials from the gradients acquired in the known area are detectable in the unknown area. In the case of the Ost and Nym areas, we can expect almost the same material composition. This should be also the case for different cities in Germany that consist of similar urban neighborhoods. Future studies shall be dedicated to test the transferability of gradients from one city to a similar but different city. To build a more robust model for several cities, gradients can also be derived from test areas of different cities with varying surface material compositions.

Although the results confirm that urban material gradients are transferable between two study sites, the physical significance of the gradients produced by PCA changed slightly between the Ost to Nym ordination spaces. In the Ost ordination space, the negative end of PC1 represents vegetation classes (including deciduous tree, lawn, and meadow), and the positive end of PC1 represents an abundance of artificial materials (e.g., cobblestone, asphalt, roofing tar, and concrete). Thus, the negative end of PC2 can be used to discriminate deciduous trees and meadow. In the Nym ordination space, PC1 can still differentiate the vegetation species and artificial materials but in a less distinctive manner; e.g., meadow is not clearly separated by PC1 any more. In addition, PC2 cannot be used to quantitatively distinguish deciduous trees and lawn. However, it should be noted here that PCA is not the optimal method for the interpretation of physical significance of urban gradients (see Ji

et al., 2020) but rather is designed to test the transferability that was the main objective of this paper. Therefore, we suggest that other ordination methods should be used to obtain the most meaningful urban gradients such as shown in (Jilge et al., 2019). However, PCA is still considered as one of the most appropriate methods for the transferablity analysis of urban gradients across different areas because it can easily transfer gradients with loadings between different study areas. Moreover, the resulting urban material gradients obtained from different ordination methods usually have similar properties. Thus, since we demonstrate the transferability of the gradients determined by PCA, the gradients determined by other ordination methods are transferable in the same situation.

The prediction maps are comparable to those acquired by (Jilge et al., 2019). Although (Jilge et al., 2019) applied detrended correspondence analysis (DCA), the prediction maps obtained in their study contain similar information to that obtained in this work, in particular for PC1. Considering these results together with the above discussion suggests that ordination methods can affect the prediction results, but will not change them completely. (Skowronek et al., 2018) et al. evaluated the transferability of HSI-based distribution models for the detection of an invasive alien bryophyte. (Skowronek et al., 2018) concluded that the success of transfer models calibrated in one site to another site depend strongly on the respective study sites. Two or more ordination methods are suggested to be used in parallel to enhance the detection of artefacts in the results, because each ordination methods with the different weightings of the elements in the species abundance matrix, and thus may explain the observed variation in the analysed ordination results. It will be interesting to examine these approaches in future studies in order to enhance our understanding of the functionality, robustness and feasibility of the methods for deriving urban gradients.

4.5.3 Potential applications of transferable urban gradients

Transferable urban material gradients can be used for time- and cost-efficient large-scale mapping of urban materials. The potential use of remote sensing images for urban mapping has studied extensively over the past decade (Ridd, 1995; Weng, 2012). Since spaceborne HSIs cover large geographical areas in high geometric detail and with a short revisiting time, their capabilities were demonstrated. However, some of the urgently needed detailed information cannot be obtained from HSIs and must be derived from other sources. To create classification maps that are useful for urban planners, supervised classification methods are commonly implemented on HSIs. These rules lead to results with an accuracy that is strongly influenced by the amount of training data. Obtaining appropriate ground truth data for implementation and validation purposes requires intense efforts in terms of time consumption and economic resources. For most areas, in situ data are either completely absent or are

outdated and unreliable. Therefore, the limited availability of in situ data is a challenge for classification problems, particularly with regard to the model transferability. The transferred urban gradients provides a possible approach for avoiding training data collection in the area that has an known area in close proximity and fits the transferable urban gradients.

Transferable urban gradients provide a practical method to obtain a fuzzy map of an unknown area with limited information and therefore can be used to improve the results of other urban mapping models. First, the transfer of urban gradients can be an useful approach for mapping urban materials when limited resources are available to carry out fieldwork and remote sensing data are available for a larger area. With a limited training set, classification accuracy tends to decrease as the number of features increases which is known as the Hughes effect (Hughes, 1968). As an increasing number of mathematical or machine learning methods are proposed with the requirement of sufficient prior knowledge, transferable urban material gradients can provide more knowledge-based information for use in these algorithms. The prior knowledge including spatial relationships and patterns of urban structures can be used to improve the characterization of not only single pixels but also of the whole image (Plaza et al., 2009).

4.6 Conclusion

Gradient analysis has the potential to be applicable to images from the ongoing and future spaceborne imaging spectroscopy missions. Although the spatial resolution of these data is considered to be coarse for urban applications and urban object-related information cannot be directly detected, it enables the derivation of surface material compositions of large areas, which is important information for continental to global urban climate related analyses.

In this paper, we addressed the question of whether gradient analysis can be a robust and transferable technique despite its data-driven nature. For this purpose, we designed three tests for simulating the transferability of urban material gradients to the Ostbahnhof area in Munich, Germany.

In the first step, we evaluated the similarity of the sample distributions in two ordination spaces, one built by the samples of the Ostbahnhof area and the other generated by the samples of the Nymphenburg area. Both gradient spaces are highly comparable, providing an initial indication of the robustness of the urban gradients in the case where the overall surface material composition is similar. It can be assumed that these gradients are applicable to other cities with similar urban structures and thus surface material compositions, so that this method will be valid for a wide range of mid-European cities. However, if new and

region-specific materials are dominating the surface material composition such as for cities with other urban structures, the gradients may differ.

We expanded the transferability test to regress the gradient scores against the surface material reflectances using PLSR and applied the resulting models to predict the surface material compositions of Ostbahnhof area using imaging spectroscopy data. The comparison of the prediction results of approach-OstOst and approach-OstNym demonstrates that the material gradients acquired from the Nym area can successfully interpret the Ost area, while the comparison of approach-OstOst and approach-NymNym show that the PLSR model retrieved from Nym area cannot be simply transferred to the Ost area.

Since this contradicts the results of the gradient space analyses, we found that the reflectance data of the two investigated areas have significant differences in the albedo despite their similar surface material compositions. This can be related to the different flight lines of the source airborne data from HyMap used for the simulation of spaceborne EnMAP data. Although we cannot fully prove the transferability of the PLSR models to different areas, the results indicate the transferability potential if well-calibrated spaceborne imaging spectroscopy data are used. Moreover, these results reveal the importance of calibrated spaceborne imaging spectroscopy data and data cross-calibration, if different spaceborne sensor data are combined.

Transferable urban material gradients can be used effectively in time-consuming and costly large-scale mapping of urban material compositions. Furthermore, they can provide a fuzzy map of an unknown area with limited information and therefore can be used to enhance the results of other urban mapping models. Although the gradient concept works well in ecology for mixed vegetation, the ability of this approach should be further tested in the field of urban material composition. Exploration of urban material gradients, focusing not only on its transferability but also on pattern recognition capability, will provide us with a more accurate and definite answer to this question.

Acknowledgment

This work was supported by the China Scholarship Council under scholarship No. 201806220088. We would like to thank Dr. Wieke Heldens and Dr. Marianne Jilge from the German Aerospace Center (DLR) for providing and pre-processing the surface material map of Munich.

Chapter 5

Synthesis

This work focuses on mapping urban surface materials using imaging spectroscopy (IS) data at different spatial scales. In particular, this study attempts to explore the information of these data to identify the chemical and physical properties of surface materials and solve the mixed pixel problem. In the following sections, the findings are summarized (subsection 5.1.1) and discussed (subsection 5.1.2) with respect to the research objectives and in the broader context. Section 5.2 highlights further investigations and experimentation that can be performed from both technical and application perspectives. Finally, Section 5.3 then presents the conclusions.

5.1 Summary and discussion

5.1.1 Summary

Mapping urban surface material is not straightforward due to the complexity of surface materials themselves and the heterogeneity of their distribution. IS data provides narrow and continuous inspection in the spectral dimension, and can therefore be beneficial in urban surface material mapping. This work begins with the detection of solar photovoltaic (PV) using airborne IS data (Chapter 2). Six spectral indices were proposed and combined, the spectral variance due to different detection angles and installation angles of solar PV systems was solved. The validation of the proposed spectral feature-based approach using the large airborne spectral library proved the ability of this approach to identify PV material among other complex urban surface materials. And the final validation in Oldenburg, Germany, demonstrates that both large PV systems on solar power plants and small PV modules on rooftops are accurately detected, further highlighting the applicability of the approach.

Due to the limitations of airborne IS data, such as low coverage, lack of time contiguity, and high cost, the thesis then focuses on urban surface material mapping with spaceborne IS data, i.e. simulated EnMAP data. The EnMAP data overcomes the limitations of airborne data but suffers from the mixed pixel problem resulting from the 30 m spatial resolution. Based on the urban gradient concept proposed by Jilge et al. (2019), the mixed pixel problem can be avoided by treating all pixels as mixed and without the requirement of pure pixels as input. However, the urban surface material gradients were generated using a data-driven method, i.e., dimension reduction, and therefore the sampling robustness and area transferability of this technique need to be analyzed before the wide usage. In this case, Chapter 3 describes how six systematic sampling schemes and three random sampling schemes were established to generate urban surface material gradients. Results show that the gradients generated from six slightly shifted systematic sampling schemes and three random sampling schemes are very similar to each other, and the distributions of the samples in the gradient spaces are also similar, as evidenced by Mantel statistical test. Therefore, the study proved the sampling robustness of urban surface material gradients.

In Chapter 4, two subsets in Munich city were employed, referred to as Ost and Nym areas. Three gradient analysis approaches were designed to prove the transferability of urban surface material gradients between Ost and Nym areas for the cases where (i) sufficient samples were collected in the Ost area; (ii) some samples were collected in the Ost area; and (iii) no samples were collected in the Ost area. The Ost samples were used to generate a gradient ordination space in case (i), while the Nym samples were used to create the ordination space to support the pattern recognition of the Ost area in cases (ii) and (iii). The findings show that the area transferability between similar urban areas was valid. Furthermore, this study proposes a complete chain to compare the data transferability of the reduction approach. It can be used not only for the applications for the transferability of gradient concept but also for other data-driven pattern description approaches in general.

5.1.2 Discussion

The thesis firstly explores the physical absorption and reflectance features of PV modules and developed a robust approach for PV modules detection. The included image spectral library takes into account all other common urban surface materials, builds the basic ground for eliminating the confusion of PV and non-PV materials. The laboratory goniometer was also considered, which monitors the condition with different illumination angles and detection angles that could happen in different PV installed rooftops and solar irradiation angles within days and years. The HySpex airborne data presents the detection noise in a real case. Considering all noise integrated enables this spectral feature-based approach a robust and

transferable approach, can serve as highly accurate PV detection in different regions. Since the thresholds of the spectral indices were empirically derived through a manual selection based on spectral libraries, the transferability of the approach may be limited in different contexts, e.g., with different sensors or different urban environments. The employed spectral libraries include goniometer data, HyMap spectral libraries, and HySpex imagery. In this way, thresholds are more generalized and not only optimized to one dataset, thereby reducing the transferability limitation. Another restriction of this approach is that thin-film panels cannot be detected because they use different materials than EVA-coated C-Si PV modules. As aforementioned, the thin-film panels are very rarely present, so this would not significantly affect the efficiency of our approach. If this PV panel material becomes more dominant, the approach needs to be extended to consider the characteristic spectral features of the material. In addition, very small PV modules also may be missed due to the relatively coarse spatial resolution of HySpex data (1.2 m) compared to the spatial occurrence of these PV panels. Therefore, other data sources that could improve the spatial resolution should be taken into account in the future.

The analysis of sampling robustness of urban gradients aims to address the effectiveness of different locations of samples to urban material gradients, and the proposed experiments have sufficiently demonstrated this in a reliable way. Because with similar configuration settings and same test sites, the size and number of samples can be copied, but the location of samples may vary from different experts' operations. We chose principle component analysis (PCA) to extract material gradients, as it is based on a linear distribution model and allows for better comparability of different material gradients. However, the PCA is not the most appropriate ordination method regarding its physical significance. It is used to focus on the first important information to address the sampling robustness problem. The extraction of material gradients is basically a dimensionality reduction of a sample-by-material matrix, and thus, can be performed by many methods. Although the resulting material gradients show generally similar characteristics, the available methods differ in their principles, performance, and fine details of the extracted material gradients. This study aims to prove the sampling robustness of urban gradient analysis. Thus, we used PCA because it is the most commonly used ordination method, and PCA provides loadings that can be used to easily and repeatedly transform samples from a material composition matrix to a gradient score matrix. Therefore, PCA is good enough for this study, and testing of other ordination methods can be pursued in the future. Another limitation is that a possible time difference between the acquisition time of the material map and IS data may occur, which could affect the gradient mapping results. In this case, there should be an uncertainty assessment with the final gradient maps.

In the analysis of the transferability of urban gradients to other areas, the generic technique of gradients is further investigated. It evaluates the direct effect, similarity, and difference in terms of transferability of the gradient approach to other urban areas despite their data-driven nature. This is very important for the future application of spaceborne IS data covering urban areas to quantitatively characterize urban spaces. Theoretically, the transferability of gradients from East to Nym is successful because it reveals the most important gradient changes. However, the three approaches also show that although the PLSR value is high, the actual application cannot describe the detailed gradual change. It is suggested to use two or more ordination methods in parallel to improve the detection of artifacts in the results since each ordination method has different weights of the elements in the species abundance matrix and thus can explain the observed variation in the analyzed ordination results. Therefore, the success of transfer models calibrated from one site to another is highly dependent on the particular study sites. The cities with similar structures and material compositions can be treated as a similar city class and therefore the transferable gradients generated by a city or cities within that city class can be readily applied. This leads to the general conclusion that the general trend can be retrieved from multiple locations, regionally or globally. In addition, in situ data is also required as a dependent input or complementary information for a detailed and comprehensive understanding or study of a city/area.

5.2 Outlook

Overall, this thesis provides very promising results, showing that IS data can contribute greatly to urban surface materials mapping. Future work can build on these results and can enhance the developed and applied methods. Two aspects, including PV detection and gradient analysis using IS data, are presented in subsection 5.2.1 and subsection 5.2.2 with their future directions.

5.2.1 Solar panel detection

From the technical perspective, in order to further improve the thresholds towards general use, advanced techniques such as machine learning can be applied, and additional datasets can also be included. The use of machine learning methods to define the thresholds for each spectral index based on training with massive pure spectra can be further developed in the next step. In addition, adding the texture, shape or other visual features from high spatial resolution data is of great potential for future studies. This study focuses on exploring the full spectral potential of IS data, which is shown to achieve very good quality with a high

potential for wider applicability due to its physics-based principles. Other studies of PV detection using machine learning methods proved that the texture or shape information with high spatial resolution color images could be independently used for PV detection. Combined with shape- and texture-based methods would greatly expand the possibility to identify PV panels and additionally reconstruct their spatial shape given that very high spatial resolution data is used. Including the texture or shape information or other visual features apart from spectral features theoretically could be helpful for the overall detection accuracy, especially for the shapes of the panels. From the application perspective, although thin-film panels are rarely used today if this PV panel material becomes more dominant, the approach shall be expanded to find characteristic spectral features of the material.

5.2.2 Urban surface material gradients

The two gradient analyses within this thesis prove the robustness and transferability of this technique for determining urban surface material compositions using upcoming spaceborne IS data. Since the mixed pixel problem complicates the use of these data for urban object detection, urban surface material gradients provide an alternative way to circumvent this problem. This new technique can be further improved or analyzed before widespread application.

Future technical work on the robustness and transferability of urban material gradients can be further analyzed. Regarding the sampling robustness, both the size of samples and the distance between samples may affect the material gradient and need to be further investigated. For the area transferability of urban surface material gradients, other potential factors may impact, such as ordination methods, detailed ground truth data, and the heterogeneity of urban areas, which need to be investigated in future work.

From the application perspective, we can use the results of the urban surface material gradient for the analysis of spatial autocorrelation in an urban environment. Tobler's law provides the theoretical basis for urban neighborhoods, which is called spatial autocorrelation but has not been proven with urban surface materials. Spatial autocorrelation, i.e., spatial association, or spatial dependence, results primarily from community processes in urban areas or from the physical forcing of environmental variables in vegetation ecology (Legendre, 1993). Finding out whether spatial autocorrelation occurs in urban surface materials, and if so, how it affects the materials, would be of great importance for urban model study. The information on the spatial patterns of surface materials would strongly support the urban neighborhood and provide the reference scale for urban functional units.

5.3 Conclusion

From a general perspective, the overarching goal of this thesis was to develop, apply, and evaluate tailored methods and approaches that implement urban surface material mapping based on IS data. This thesis provides very promising results, which show that IS data has a high capability to contribute to detailed and specific urban surface material detection on a local or regional scale, and global scale with upcoming spaceborne spectrometers. The first novelty of this work presented in Chapter 2 is developing and combining spectral indices based on large spectral libraries, which can account for inter- and intra- class spectral variability for PV modules detection. This work on the one side creates greater awareness of the potential importance and applicability of IS data for PV modules identification, on the other side advances the use of IS data-based application on general material identification. Secondly, when turning to regional or global material detection, employed spaceborne IS data meet critical mixed pixel problems, and therefore the urban surface material gradient analysis was established based on the existence of urban neighborhoods. The sampling robustness evaluation of the gradients (presented in Chapter 3) not only strengthens the gradient approach from the sampling strategies but also creates a generic way to evaluate the sampling robustness in similar circumstances. The third novelty of this work (presented in Chapter 4) is the area transferability analysis of urban surface material gradients, which is the most concerned question in the application of new methods. The analyses of sampling robustness and area transferability pave the way for subsequent analysis of the stability of urban surface material gradients and the interpretation of material gradients in other urban environments.

Therefore, the proposed methods in this thesis contribute to a more detailed understanding of urban surface material mapping. This comprehensive understanding is essential for urban climate models and related urban planning by local authorities. In addition, this thesis attempts to step back and understand the key factors in urban surface mapping using IS data from the perspective of source data characteristics. Among three essential factors, i.e. spectral resolution, spatial resolution, and detection repeatability, spectral resolution should be considered primarily, which is why IS data is the most suitable choice for urban surface mapping. Depending on the objectives, the high spatial resolution and high repeatability are trade-offs which differ between airborne IS data and spaceborne IS data. From the methodological perspective, this thesis attempted to categorize various state-of-the-art methods that benefit from multiple data sources, classification units, and mapping patterns.

As Big Data and global mapping are popular and point to the future, space-based IS data is potentially more promising on mapping urban surfaces. Since the spatial resolution of space-based IS data is relatively coarse, pixels are often heavily mixed with a combination

of materials. In this context, assigning spectra from pixels on IS data to fixed classes of surface materials is not accurate enough, and information could easily be lost. In this context, gradient mapping could best preserve the information in the original data and does not require a prior determination of pure pixels, which most pixel unmixing methods have problems with space-based IS data. Therefore, it would be beneficial to combine hard classification and gradient mapping to improve the classification capability in urban mapping in the future.

References

T. Acuña-Ruz, D. Uribe, R. Taylor, L. Amézquita, M. C. Guzmán, J. Merrill, P. Martínez, L. Voisin, and C. Mattar. Anthropogenic marine debris over beaches: Spectral characterization for remote sensing applications. *Remote Sensing of Environment*, 217:309–322, 2018.

M. Alberti. The effects of urban patterns on ecosystem function. *International regional science review*, 28(2):168–192, 2005.

K. Alonso, M. Bachmann, K. Burch, E. Carmona, D. Cerra, R. de los Reyes, D. Dietrich, U. Heiden, A. Hölderlin, J. Ickes, U. Knodt, D. Krutz, H. Lester, R. Müller, M. Pagnutti, P. Reinartz, R. Richter, R. Ryan, I. Sebastian, and M. Tegler. Data Products, Quality and Validation of the DLR Earth Sensing Imaging Spectrometer (DESIS). *Sensors*, 19(20): 4471, Jan. 2019.

A. H. Auer Jr. Correlation of land use and cover with meteorological anomalies. *Journal of Applied Meteorology*, 17(5):636–643, 1978.

K. Austrian, J. Pinchoff, J. B. Tidwell, C. White, T. Abuya, B. Kangwana, R. Ochako, J. Wanyungu, E. Muluve, F. Mbushi, et al. Covid-19 related knowledge, attitudes, practices and needs of households in informal settlements in nairobi, kenya. 2020.

M. Bachmann, K. Alonso, E. Carmona, B. Gerasch, M. Habermeyer, S. Holzwarth, H. Krawczyk, M. Langheinrich, D. Marshall, M. Pato, et al. Analysis-ready data from hyperspectral sensors—the design of the enmap card4l-sr data product. *Remote Sensing*, 13(22):4536, 2021.

E. Ben-Dor and F. Kruse. Surface mineral mapping of makhtesh ramon negev, israel using ger 63 channel scanner data. *International Journal of Remote Sensing*, 16(18):3529–3553, 1995.

J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE journal of selected topics in applied earth observations and remote sensing*, 5(2):354–379, 2012.

R. Bivand, T. Keitt, B. Rowlingson, E. Pebesma, M. Sumner, R. Hijmans, E. Rouault, and M. R. Bivand. Package 'rgdal'. *Bindings for the Geospatial Data Abstraction Library. Available online: https://cran. r-project. org/web/packages/rgdal/index. html (accessed on 15 October 2017)*, 2015.

M. Bochow. *Automatisierungspotenzial von Stadtbiotopkartierungen durch Methoden der Fernerkundung*. Logos Verlag Berlin GmbH, 2010.

M. Bochow, K. Segl, and H. Kaufmann. An update system for urban biotope maps based on hyperspectral remote sensing data. In *Proceedings 5th EARSeL Workshop on Imaging Spectroscopy*, volume 1, Bruges, 2007.

J. Broadwater and A. Banerjee. A comparison of kernel functions for intimate mixture models. In *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pages 1–4. IEEE, 2009.

B.-G. Burduhos, I. Vişa, A. Duţă, and M. Neagoe. Analysis of the conversion efficiency of five types of photovoltaic modules during high relative humidity time periods. *IEEE Journal of Photovoltaics*, 8(6):1716–1724, 2018.

J. Camilo, R. Wang, L. M. Collins, K. Bradbury, and J. M. Malof. Application of a semantic segmentation convolutional neural network for accurate automatic detection and mapping of solar photovoltaic arrays in aerial imagery. *arXiv preprint:1801.04018*, 2018.

J. B. Campbell and R. H. Wynne. *Introduction to remote sensing*. Guilford Press, 2011.

W. Cao, L. Dong, L. Wu, and Y. Liu. Quantifying urban areas with multi-source data based on percolation theory. *Remote Sensing of Environment*, 241:111730, May 2020. ISSN 0034-4257.

F. Castaldi, S. Chabrillat, and B. Van Wesemael. Sampling strategies for soil property mapping using multispectral sentinel-2 and hyperspectral enmap satellite data. *Remote Sensing*, 11(3):309, 2019.

K. Cawse-Nicholson, P. A. Townsend, D. Schimel, A. M. Assiri, P. L. Blake, M. F. Buongiorno, P. Campbell, N. Carmon, K. A. Casey, R. E. Correa-Pabón, et al. Nasa's surface biology and geology designated observable: A perspective on surface imaging algorithms. *Remote Sensing of Environment*, 257:112349, 2021.

F. Chen, R. Bornstein, S. Grimmond, J. Li, X. Liang, A. Martilli, S. Miao, J. Voogt, and Y. Wang. Research priorities in observing and modeling urban weather and climate. *Bull. Am. Meteorol. Soc.*, 93:1725–1728, 2012.

F. Chen, H. Jiang, T. Van de Voorde, S. Lu, W. Xu, and Y. Zhou. Land cover mapping in urban environments using hyperspectral apex data: A study case in baden, switzerland. *International journal of applied earth observation and geoinformation*, 71:70–82, 2018.

J. Chen and J. Chen. Globeland30: Operational global land cover mapping and big-data analysis. *Science China. Earth Sciences*, 61(10):1533–1534, 2018.

J. Chen, J. Chen, A. Liao, X. Cao, L. Chen, X. Chen, C. He, G. Han, S. Peng, M. Lu, et al. Global land cover mapping at 30 m resolution: A pok-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103:7–27, 2015.

R. N. Clark and T. L. Roush. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, 89(B7): 6329–6340, 1984.

R. N. Clark, T. V. King, M. Klejwa, G. A. Swayze, and N. Vergo. High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research: Solid Earth*, 95 (B8):12653–12680, 1990.

R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent. Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems. *Journal of Geophysical Research: Planets*, 108(E12), 2003.

E. A. Cloutis. Spectral reflectance properties of hydrocarbons: remote-sensing implications. *Science*, 245(4914):165–168, 1989.

E. A. Cloutis. Review article hyperspectral geological remote sensing: evaluation of analytical techniques. *International Journal of Remote Sensing*, 17(12):2215–2242, 1996.

T. Cocks, R. Jenssen, A. Stewart, I. Wilson, and T. Shields. The hymap airborne hyperspectral sensor: the system, calibration and performance. In *Proceedings of the 1st EARSeL Workshop on Imaging Spectroscopy*, pages 37–42, Zurich, Swiss, 1998.

S. Cogliati, F. Sarti, L. Chiarantini, M. Cosi, R. Lorusso, E. Lopinto, F. Miglietta, L. Genesio, L. Guanter, A. Damm, et al. The prisma imaging spectroscopy mission: overview and first performance analysis. *Remote Sensing of Environment*, 262:112499, 2021.

S. Cooper, A. Okujeni, D. Pflugmacher, S. van der Linden, and P. Hostert. Combining simulated hyperspectral enmap and landsat time series for forest aboveground biomass mapping. *International Journal of Applied Earth Observation and Geoinformation*, 98: 102307, 2021.

C. Corbane, M. Pesaresi, T. Kemper, P. Politis, A. J. Florczyk, V. Syrris, M. Melchiorri, F. Sabo, and P. Soille. Automated global delineation of human settlements from 40 years of landsat satellite data archives. *Big Earth Data*, 3(2):140–169, 2019.

J. Corburn, D. Vlahov, B. Mberu, L. Riley, W. T. Caiaffa, S. F. Rashid, A. Ko, S. Patel, S. Jukur, E. Martínez-Herrera, et al. Slum health: arresting covid-19 and improving well-being in urban informal settlements. *Journal of urban health*, 97(3):348–357, 2020.

B. Cui, Q. Zhao, W. Huang, X. Song, H. Ye, and X. Zhou. Leaf chlorophyll content retrieval of wheat by simulated rapideye, sentinel-2 and enmap data. *Journal of Integrative Agriculture*, 18(6):1230–1245, 2019.

D. Czirjak. Detecting photovoltaic solar panels using hyperspectral imagery and estimating solar power production. *Journal of Applied Remote Sensing*, 11(2):026007, Apr. 2017. ISSN 1931-3195.

S. M. Davis, D. A. Landgrebe, T. L. Phillips, P. H. Swain, R. M. Hoffer, J. C. Lindenlaub, and L. F. Silva. Remote sensing: the quantitative approach. *New York*, 1978.

J. de Hoog, S. Maetschke, P. Ilfrich, and R. R. Kolluri. Using satellite and aerial imagery for identification of solar pv: State of the art and research opportunities. In *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, pages 308–313, 2020.

J. Degerickx, D. A. Roberts, J. P. McFadden, M. Hermy, and B. Somers. Urban tree health assessment using airborne hyperspectral and lidar imagery. *International journal of applied earth observation and geoinformation*, 73:26–38, 2018.

Q. Deng, Z. Wang, S. Wang, and G. Shao. Simulation of planar si/mg2si/si pin heterojunction solar cells for high efficiency. *Solar Energy*, 158:654–662, 2017.

U. DESA. 2018 revision of world urbanization prospects, 2018.

R. Dian, S. Li, B. Sun, and A. Guo. Recent advances and new guidelines on hyperspectral and multispectral image fusion. *Information Fusion*, 2020.

P. Dixon. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, 14(6):927–930, 2003. ISSN 1654-1103.

N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero. Nonlinear unmixing of hyperspectral images: Models and algorithms. *IEEE Signal Processing Magazine*, 31(1):82–94, 2013.

R. Dunn and A. R. Harrison. Two-dimensional systematic sampling of land use. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 42(4):585–601, 1993. ISSN 1467-9876.

I. Eliasson. The use of climate knowledge in urban planning. *Landscape and urban planning*, 48(1-2):31–44, 2000.

T. Esch, M. Marconcini, A. Felbier, A. Roth, W. Heldens, M. Huber, M. Schwinger, H. Taubenböck, A. Müller, and S. Dech. Urban footprint processor—fully automated processing chain generating settlement masks from global data of the tandem-x mission. *IEEE Geoscience and Remote Sensing Letters*, 10(6):1617–1621, 2013.

T. Esch, W. Heldens, A. Hirner, M. Keil, M. Marconcini, A. Roth, J. Zeidler, S. Dech, and E. Strano. Breaking new ground in mapping human settlements from space–the global urban footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134:30–42, 2017.

T. Esch, E. Brzoska, S. Dech, B. Leutner, D. Palacios-Lopez, A. Metz-Marconcini, M. Marconcini, A. Roth, and J. Zeidler. World settlement footprint 3d-a first three-dimensional survey of the global building stock. *Remote Sensing of Environment*, 270:112877, 2022.

M. Eslami and A. Mohammadzadeh. Developing a spectral-based strategy for urban object detection from airborne hyperspectral tir and visible data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5):1808–1816, 2015.

H. Feilhauer and S. Schmidtlein. Mapping continuous fields of forest alpha and beta diversity. *Applied Vegetation Science*, 12(4):429–439, 2009. ISSN 1654-109X.

H. Feilhauer, U. Faude, and S. Schmidtlein. Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. *Remote Sensing of Environment*, 115(10):2513–2524, Oct. 2011. ISSN 00344257.

H. Feilhauer, C. Dahlke, D. Doktor, A. Lausch, S. Schmidtlein, G. Schulz, and S. Stenzel. Mapping the local variability of natura 2000 habitats with remote sensing. *Applied vegetation science*, 17(4):765–779, 2014.

H. Feilhauer, A. Zlinszky, A. Kania, G. M. Foody, D. Doktor, A. Lausch, and S. Schmidtlein. Let your maps be fuzzy!—class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sensing in Ecology and Conservation*, 2020.

B. Feizizadeh, D. Omarzadeh, Z. Ronagh, A. Sharifi, T. Blaschke, and T. Lakes. A scenario-based approach for urban water management in the context of the covid-19 pandemic and a case study for the tabriz metropolitan area, iran. *Science of The Total Environment*, 790: 148272, 2021.

J. Fenger. Urban air quality. *Atmospheric environment*, 33(29):4877–4900, 1999.

J. Franke, D. A. Roberts, K. Halligan, and G. Menz. Hierarchical multiple endmember spectral mixture analysis (mesma) of hyperspectral imagery for urban environments. *Remote Sensing of Environment*, 113(8):1712–1723, 2009.

S. Frontier. Decrease of eigenvalues in principal component analysis-comparison with broken stick model. *Journal of Experimental Marine Biology and Ecology*, 25(1):67–75, Nov. 1976. ISSN 0022-0981.

K. P. F.R.S. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, Nov. 1901. ISSN 1941-5982.

S. P. Garaba and H. M. Dierssen. An airborne remote sensing case study of synthetic hydrocarbon detection using short wave infrared absorption features identified from marine-harvested macro-and microplastics. *Remote Sensing of Environment*, 205:224–235, 2018.

H. A. Gleason. The individualistic concept of the plant association. *Bulletin of the Torrey botanical club*, pages 7–26, 1926.

A. A. Green, M. Berman, P. Switzer, and M. D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, 26(1):65–74, Jan. 1988. ISSN 0196-2892.

M. A. Green. How did solar cells get so cheap? *Joule*, 3(3):631–633, 2019.

D. A. Griffith. Effective geographic sample size in the presence of spatial autocorrelation. *Annals of the Association of American Geographers*, 95(4):740–760, 2005.

N. B. Grimm, S. H. Faeth, N. E. Golubiewski, C. L. Redman, J. Wu, X. Bai, and J. M. Briggs. Global change and the ecology of cities. *science*, 319(5864):756–760, 2008.

R. Grote, R. Samson, R. Alonso, J. H. Amorim, P. Cariñanos, G. Churkina, S. Fares, D. L. Thiec, Ü. Niinemets, T. N. Mikkelsen, et al. Functional traits of urban trees: air pollution mitigation potential. *Frontiers in Ecology and the Environment*, 14(10):543–550, 2016.

H. Gu, A. Singh, and P. A. Townsend. Detection of gradients of forest composition in an urban area using imaging spectroscopy. *Remote Sensing of Environment*, 167:168–180, 2015.

L. Guanter, H. Kaufmann, K. Segl, S. Foerster, C. Rogass, S. Chabrillat, T. Kuester, A. Hollstein, G. Rossner, C. Chlebek, C. Straif, S. Fischer, S. Schrader, T. Storch, U. Heiden, A. Mueller, M. Bachmann, H. Mühle, R. Müller, M. Habermeyer, A. Ohndorf, J. Hill, H. Buddenbaum, P. Hostert, S. Van der Linden, P. J. Leitão, A. Rabe, R. Doerffer, H. Krasemann, H. Xi, W. Mauser, T. Hank, M. Locherer, M. Rast, K. Staenz, and B. Sang. The EnMAP spaceborne imaging spectroscopy mission for earth observation. *Remote Sensing*, 7(7):8830–8857, 2015.

X. Guo and P. Li. Mapping plastic materials in an urban area: Development of the normalized difference plastic index using worldview-3 superspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169:214–226, 2020.

M. Habermeyer, U. Marschalk, and A. Roth. Digital elevation model database W42 - a scalable system for spatial data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B1):1253–1258, July 2008.

U. Habitat. World cities report 2020: The value of sustainable urbanization, 2020.

U. Heiden. *Analyse hyperspektraler Flugzeugscannerdaten zur ökologischen Charakterisierung städtischer Biotope*. PhD thesis, Dissertation, Technische Universität Berlin, 2004.

U. Heiden, K. Segl, S. Roessner, and H. Kaufmann. Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data. *Remote Sensing of Environment*, 111(4):537–552, 2007.

U. Heiden, W. Heldens, S. Roessner, K. Segl, T. Esch, and A. Mueller. Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and urban Planning*, 105(4):361–375, 2012.

W. Heldens. *Use of airborne hyperspectral data and height information to support urban micro climate characterisation*. PhD thesis, University Wuerzburg, Germany, 2010.

M. Herold and D. Roberts. Spectral characteristics of asphalt road aging and deterioration: implications for remote-sensing applications. *Applied optics*, 44(20):4327–4334, 2005.

M. Herold, J. Scepan, and K. C. Clarke. The use of remote sensing and landscape metrics to describe structures and changes in urban land uses. *Environment and planning A*, 34(8): 1443–1458, 2002.

M. Herold, M. E. Gardner, and D. A. Roberts. Spectral resolution requirements for mapping urban areas. *IEEE Transactions on Geoscience and remote sensing*, 41(9):1907–1919, 2003.

M. Herold, D. A. Roberts, M. E. Gardner, and P. E. Dennison. Spectrometry for urban area remote sensing—development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, 91(3-4):304–319, 2004.

B. W. Heumann. Satellite remote sensing of mangrove forests: Recent advances and future opportunities. *Progress in Physical Geography*, 35(1):87–108, 2011.

R. J. Hijmans, J. van Etten, M. Mattiuzzi, M. Sumner, J. Greenberg, O. Lamigueiro, A. Bevan, E. Racine, and A. Shortridge. Raster package in r, 2013.

M. O. Hill and H. G. Gauch. *Detrended correspondence analysis: an improved ordination technique*, volume 2. Springer, Dordrecht, advances in vegetation science edition, 1980. ISBN 978-94-009-9199-6.

B. Hörig, F. Kühn, F. Oschütz, and F. Lehmann. Hymap hyperspectral remote sensing to detect hydrocarbons. *International Journal of Remote Sensing*, 22(8):1413–1422, 2001.

H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441, 1933. ISSN 1939-2176(Electronic),0022-0663(Print).

G. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1):55–63, Jan. 1968. ISSN 1557-9654. Conference Name: IEEE Transactions on Information Theory.

G. R. Hunt and J. W. Salisbury. Assessment of landsat filters for rock type discrimination, based on intrinsic information in laboratory spectra. *Geophysics*, 43(4):738–747, 1978.

D. A. Jackson. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology*, 74(8), 1993. ISSN 1939-9170.

C. Ji, M. Jilge, U. Heiden, M. Stellmes, and H. Feilhauer. Sampling robustness in gradient analysis of urban material mixtures. *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

C. Ji, M. Bachmann, T. Esch, H. Feilhauer, U. Heiden, W. Heldens, A. Hueni, T. Lakes, A. Metz-Marconcini, M. Schroedter-Homscheidt, et al. Solar photovoltaic module detection using laboratory and airborne imaging spectroscopy data. *Remote Sensing of Environment*, 266:112692, 2021a.

C. Ji, D. Cerra, E. Carmona, S. Wilbert, N. Hanrieder, and U. Heiden. Potential of desis data to characterize solar photovoltaic (pv) systems. In *the 1st DESIS User Workshop*, 2021b.

C. Ji, U. Heiden, T. Lakes, and H. Feilhauer. Are urban material gradients transferable between areas? *International Journal of Applied Earth Observation and Geoinformation*, 100:102332, 2021c.

J. Jiang, J. Ma, Z. Wang, C. Chen, and X. Liu. Hyperspectral image classification in the presence of noisy labels. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 851–865, Feb. 2019. ISSN 0196-2892, 1558-0644.

M. Jilge. *Mapping urban surface materials with imaging spectroscopy data on different spatial scales*. PhD thesis, Ruhr-Universität Bochum, 2019.

M. Jilge, U. Heiden, C. Neumann, and H. Feilhauer. Gradients in urban material composition: A new concept to map cities with spaceborne imaging spectroscopy data. *Remote Sensing of Environment*, 223:179–193, Mar. 2019. ISSN 00344257.

Jing Wang and Chein-I Chang. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6):1586–1600, June 2006. ISSN 0196-2892.

I. Jolliffe. Principal component analysis. In M. Lovric, editor, *International Encyclopedia of Statistical Science*, pages 1094–1096. Springer Berlin Heidelberg, Berlin, 2011. ISBN 978-3-642-04898-2.

N. Kabisch, P. Selsam, T. Kirsten, A. Lausch, and J. Bumberger. A multi-sensor and multi-temporal remote sensing approach to detect land cover change dynamics in heterogeneous urban landscapes. *Ecological indicators*, 99:273–282, 2019.

A. Kantzioura, P. Kosmopoulos, and S. Zoras. Urban surface temperature and microclimate measurements in thessaloniki. *Energy and buildings*, 44:63–72, 2012.

M. S. Karoui, F. Z. Benhalouche, Y. Deville, K. Djerriri, X. Briottet, T. Houet, A. Le Bris, and C. Weber. Partial Linear NMF-Based Unmixing Methods for Detection and Area Estimation of Photovoltaic Panels in Urban Hyperspectral Remote Sensing Data. *Remote Sensing*, 11(18):2164, Jan. 2019.

C. Kennedy, J. Steinberger, B. Gasson, Y. Hansen, T. Hillman, M. Havranek, D. Pataki, A. Phdungsilp, A. Ramaswami, and G. V. Mendez. Greenhouse gas emissions from global cities, 2009.

N. Keshava. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Transactions on Geoscience and remote sensing*, 42(7):1552–1565, 2004.

S. Khanna, M. J. Santos, S. L. Ustin, A. Koltunov, R. F. Kokaly, and D. A. Roberts. Detection of salt marsh vegetation stress and recovery after the deepwater horizon oil spill in barataria bay, gulf of mexico using aviris data. *PloS one*, 8(11):e78989, 2013.

T. V. King and R. N. Clark. Spectral characteristics of chlorites and mg-serpentines using high-resolution reflectance spectroscopy. *Journal of Geophysical Research: Solid Earth*, 94(B10):13997–14008, 1989.

R. G. Knox. Effects of detrending and rescaling on correspondence analysis: solution stability and accuracy. *Vegetatio*, 83(1):129–136, Oct. 1989. ISSN 1573-5052.

C. H. Köhler. Airborne imaging spectrometer hyspex. *Journal of large-scale research facilities*, 2:A93, 2016.

R. F. Kokaly. Prism: Processing routines in idl for spectroscopic measurements (installation manual and user's guide, version 1.0). Technical report, US Geological Survey, 2011.

R. F. Kokaly and A. K. Skidmore. Plant phenolics and absorption features in vegetation reflectance spectra near 1.66 μm. *International Journal of Applied Earth Observation and Geoinformation*, 43:55–83, 2015.

R. F. Kokaly, B. R. Couvillion, J. M. Holloway, D. A. Roberts, S. L. Ustin, S. H. Peterson, S. Khanna, and S. C. Piazza. Spectroscopic remote sensing of the distribution and persistence of oil from the deepwater horizon spill in barataria bay marshes. *Remote Sensing of Environment*, 129:210–230, 2013.

H. J. Kramer. *Earth observation remote sensing: survey of missions and sensors*. Springer Science & Business Media, 2012.

F. Kruse, J. Boardman, and J. Huntington. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1388–1400, June 2003. ISSN 1558-0644.

F. A. Kruse. Use of airborne imaging spectrometer data to map minerals associated with hydrothermally altered rocks in the northern grapevine mountains, nevada, and california. *Remote Sensing of Environment*, 24(1):31–51, 1988.

J. B. Kruskal. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29 (2):115–129, June 1964. ISSN 0033-3123, 1860-0980.

D. Krutz, R. Müller, U. Knodt, B. Günther, I. Walter, I. Sebastian, T. Säuberlich, R. Reulke, E. Carmona, A. Eckardt, H. Venus, C. Fischer, B. Zender, S. Arloth, M. Lieder, M. Neidhardt, U. Grote, F. Schrandt, S. Gelmi, and A. Wojtkowiak. The instrument design of the dlr earth sensing imaging spectrometer (desis). *Sensors*, 19(7):1622, Jan. 2019. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

F. Kuehn, K. Oppermann, and B. Hörig. Hydrocarbon index–an algorithm for hyperspectral detection of hydrocarbons. *International Journal of Remote Sensing*, 25(12):2467–2473, 2004.

T. Lakes and H.-O. Kim. The urban environmental indicator "biotope area ratio"—an enhanced approach to assess and manage the urban ecosystem services using high resolution remote-sensing. *Ecological Indicators*, 13(1):93–103, 2012.

T. Lakes, D. Krajzewicz, M. Hardinghaus, and S. Nieland. Scenarios of developing sustainable urban living environments and the role of mobility. In *GI Forum*, Salzburg, 2021.

A. Lausch, S. Erasmi, D. J. King, P. Magdon, and M. Heurich. Understanding forest health with remote sensing-part i—a review of spectral traits, processes and remote-sensing characteristics. *Remote Sensing*, 8(12):1029, 2016.

C. M. Lee, M. L. Cable, S. J. Hook, R. O. Green, S. L. Ustin, D. J. Mandl, and E. M. Middleton. An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sensing of Environment*, 167:6–19, Sept. 2015. ISSN 0034-4257.

P. Legendre. Spatial autocorrelation: trouble or new paradigm? *Ecology*, 74(6):1659–1673, 1993.

P. Legendre and L. Legendre. *Numerical ecology*. Elsevier, 2012.

I. Leifer, W. J. Lehr, D. Simecek-Beatty, E. Bradley, R. Clark, P. Dennison, Y. Hu, S. Matheson, C. E. Jones, B. Holt, et al. State of the art satellite and airborne marine oil spill remote sensing: Application to the bp deepwater horizon oil spill. *Remote Sensing of Environment*, 124:185–209, 2012.

N. Levin, R. Lugassi, U. Ramon, O. Braun, and E. Ben-Dor. Remote sensing as a tool for monitoring plasticulture in agricultural landscapes. *International journal of remote sensing*, 28(1):183–202, 2007.

J. Li, J. M. Bioucas-Dias, and A. Plaza. Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3):809–823, 2011.

J. Li, J. M. Bioucas-Dias, and A. Plaza. Spectral–Spatial Hyperspectral Image Segmentation Using Subspace Multinomial Logistic Regression and Markov Random Fields. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3):809–823, Mar. 2012a. ISSN 1558-0644. Conference Name: IEEE Transactions on Geoscience and Remote Sensing.

S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson. Deep learning for hyperspectral image classification: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):6690–6709, Sept. 2019. ISSN 0196-2892, 1558-0644.

W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce. Locality-preserving dimensionality reduction and classification for hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4):1185–1198, Apr. 2012b. ISSN 0196-2892.

J. Liang, J. Zhou, X. Bai, and Y. Qian. Salient object detection in hyperspectral imagery. In *2013 IEEE International Conference on Image Processing*, pages 2393–2397. IEEE, 2013.

B. Liu, Y. Li, G. Li, and A. Liu. A spectral feature based convolutional neural network for classification of sea surface oil spill. *ISPRS International Journal of Geo-Information*, 8 (4):160, 2019.

R. Loizzo, M. Daraio, R. Guarini, F. Longo, R. Lorusso, L. Dini, and E. Lopinto. Prisma Mission Status and Perspective. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 4503–4506, July 2019. doi: 10.1109/IGARSS.2019.8899272. ISSN: 2153-7003.

E. Lopinto and C. Ananasso. The Prisma hyperspectral mission. In *the 33rd EARSeL Symposium, Towards Horizon*, page 12, Matera, 2013.

S. Lu, B. Wu, N. Yan, and H. Wang. Water body mapping method with hj-1a/b satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, 13(3): 428–434, 2011.

K. Makantasis, K. Karantzalos, A. Doulamis, and K. Loupos. Deep learning-based man-made object detection from hyperspectral data. In *International symposium on visual computing*, pages 717–727. Springer, 2015.

J. M. Malof, K. Bradbury, L. M. Collins, and R. G. Newell. Automatic detection of solar photovoltaic arrays in high resolution aerial imagery. *Applied energy*, 183:229–240, 2016a.

J. M. Malof, K. Bradbury, L. M. Collins, R. G. Newell, A. Serrano, H. Wu, and S. Keene. Image features for pixel-wise detection of solar photovoltaic arrays in aerial imagery using a random forest classifier. In *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 799–803. IEEE, 2016b.

M. Marconcini, A. Metz-Marconcini, S. Üreyen, D. Palacios-Lopez, W. Hanke, F. Bachofer, J. Zeidler, T. Esch, N. Gorelick, A. Kakarla, et al. Outlining where humans live, the world settlement footprint 2015. *Scientific Data*, 7(1):1–14, 2020.

K. Mardia, J. Kent, and J. Bibby. Multivariate analysis. *Probability and mathematical statistics*, 1979.

J. Markard. The next phase of the energy transition and its implications for research and policy. *Nature Energy*, 3(8):628–633, 2018.

D. Marshall, M. Bachmann, M. Habermeyer, U. Heiden, S. Holzwarth, and T. Schmid. An automated operational processor for the determination of fractional vegetation cover from desis observations. In *the 1st DESIS User Workshop*, 2021.

J. Masek, F. Lindsay, and S. Goward. Dynamics of urban growth in the washington dc metropolitan area, 1973-1996, from landsat observations. *International Journal of Remote Sensing*, 21(18):3473–3486, 2000.

T. Matsunaga, A. Iwasaki, S. Tsuchida, K. Iwao, R. Nakamura, H. Yamamoto, S. Kato, K. Obata, O. Kashimura, J. Tanii, K. Mouri, and T. Tachikawa. HISUI status toward FY2019 launch. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 160–163, Valencia, July 2018.

T. Matsunaga, A. Iwasaki, S. Tsuchida, K. Iwao, J. Tanii, O. Kashimura, R. Nakamura, H. Yamamoto, S. Kato, K. Obata, K. Mouri, and T. Tachikawa. HISUI Status Toward 2020 Launch. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 4495–4498, July 2019. doi: 10.1109/IGARSS.2019.8899179. ISSN: 2153-7003.

R. I. McDonald, P. J. Marcotullio, and B. Güneralp. Urbanization and global trends in biodiversity and ecosystem services. In *Urbanization, biodiversity and ecosystem services: Challenges and opportunities*, pages 31–52. Springer, Dordrecht, 2013.

S. Menon, H. Akbari, S. Mahanama, I. Sednev, and R. Levinson. Radiative forcing and temperature response to changes in urban albedos and associated co2 offsets. *Environmental Research Letters*, 5(1):014005, 2010.

R. Mohanty, S. L. Happy, and A. Routray. A semisupervised spatial spectral regularized manifold local scaling cut with hgf for dimensionality reduction of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6):3423–3435, 2019. ISSN 0196-2892.

R. Mueller, M. Lehner, P. Reinartz, M. Schroeder, and R. Mueller. Evaluation of spaceborne and airborne line scanner images using a generic ortho image processor. *Proc. of High Resolution Earth Imaging for Geospatial Information, ISPRS Hannover Workshop*, 5, 2005.

C. Neumann. *Spatial nature conservation monitoring on the basis of ecological gradients using imaging spectroscopy*. PhD thesis, Technischen University Berlin, 2017. Accepted: 2017-07-18T16:10:59Z.

C. Neumann, S. Itzerott, G. Weiss, B. Kleinschmit, and S. Schmidtlein. Mapping multiple plant species abundance patterns-a multiobjective optimization procedure for combining reflectance spectroscopy and species ordination. *Ecological Informatics*, 36:61–76, 2016.

J. Nieke and M. Rast. Status: Copernicus Hyperspectral Imaging Mission For The Environment (CHIME). In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 4609–4611, July 2019. doi: 10.1109/IGARSS.2019.8899807. ISSN: 2153-7003.

J. Niemelä. Ecology and urban planning. *Biodiversity & Conservation*, 8(1):119–131, 1999.

E. Nix, J. Paulose, M. Lakhanpaul, P. Factor-Litvak, P. Parikh, H. Altamirano-Medina, Y. Bou Karim, and L. Manikam. Covid-19 & informal settlements: is stay home safe? *UCL Open: Environment Preprint*, 2021.

T. Oke. *Boundary Layer Climates*. Routledge, London, New York, 1987.

J. Oksanen, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. O'hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, H. Wagner, et al. Package 'vegan'. *Community ecology package, version*, 2(9):1–295, 2013.

A. Okujeni, S. van der Linden, and P. Hostert. Extending the vegetation–impervious–soil model using simulated enmap data and machine learning. *Remote Sensing of Environment*, 158:69–80, 2015.

D. Palacios-Lopez, F. Bachofer, T. Esch, M. Marconcini, K. MacManus, A. Sorichetta, J. Zeidler, S. Dech, A. J. Tatem, and P. Reinartz. High-resolution gridded population datasets: Exploring the capabilities of the world settlement footprint 2019 imperviousness layer for the african continent. *Remote Sensing*, 13(6):1142, 2021.

H. Pearsall and J. Pierce. Urban sustainability and environmental justice: evaluating the linkages in public planning/policy discourse. *Local Environment*, 15(6):569–580, 2010.

R. Pelta, N. Carmon, and E. Ben-Dor. A machine learning approach to detect crude oil contamination in a real scenario using hyperspectral remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 82:101901, 2019.

P. R. Peres-Neto and D. A. Jackson. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia*, 129(2):169–178, Oct. 2001. ISSN 1432-1939.

M. Pesaresi, G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. Kauffmann, T. Kemper, L. Lu, et al. A global human settlement layer from optical hr/vhr rs data: Concept and first results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(5):2102–2131, 2013.

M. Pesaresi, C. Corbane, A. Julea, A. J. Florczyk, V. Syrris, and P. Soille. Assessment of the added-value of sentinel-2 for detecting built-up areas. *Remote Sensing*, 8(4):299, 2016.

N. Pinnel, P. Gege, T. Schwarzmaier, S. Plattner, R. de los Reyes, I. A. Pérez, and C. Miller. Water quality mapping in térraba sièrpe wetland using desis and sentinel 2 data. In *the 1st DESIS User Workshop*, 2021.

A. Plaza, P. Martinez, J. Plaza, and R. Perez. Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):466–479, Mar. 2005. ISSN 0196-2892.

A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113:S110–S122, Sept. 2009. ISSN 0034-4257.

L. Plourde and R. G. Congalton. Sampling method and sample placement. *Photogrammetric Engineering & Remote Sensing*, 69(3):289–297, Mar. 2003. doi: info:doi/10.14358/PERS.69.3.289.

D. Potere, A. Schneider, S. Angel, and D. L. Civco. Mapping urban areas on a global scale: which of the eight maps now available is more accurate? *International Journal of Remote Sensing*, 30(24):6531–6558, 2009.

F. Priem and F. Canters. Synergistic use of lidar and apex hyperspectral data for high-resolution urban land cover mapping. *Remote sensing*, 8(10):787, 2016.

F. Priem, A. Okujeni, S. van der Linden, and F. Canters. Comparing map-based and library-based training approaches for urban land-cover fraction mapping from sentinel-2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, 78:295–305, 2019.

F. Priem, B. Somers, and F. Canters. Iterative spectral distancing: A novel approach for extracting endmembers in complex urban image scenes. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 4035–4038. IEEE, 2021.

QGIS Development Team. *QGIS Geographic Information System. Open Source Geospatial Foundation Project*, 2020.

D. A. Quattrochi and M. K. Ridd. Measurement and analysis of thermal energy responses from discrete urban surfaces using remote sensing data. *International Journal of Remote Sensing*, 15(10):1991–2022, 1994.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL http://www.R-project.org/.

M. Rast, J. Nieke, J. Adams, C. Isola, and F. Gascon. Copernicus hyperspectral imaging mission for the environment (chime). In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 108–111. IEEE, 2021.

R. Richter and D. Schlaepfer. Atmospheric/topographic correction for airborne imagery., 2011.

R. Richter, D. Schläpfer, and A. Müller. Operational atmospheric correction for imaging spectrometers accounting for the smile effect. *IEEE Transactions on Geoscience and Remote Sensing*, 49(5):1772–1780, 2011.

M. K. Ridd. Exploring a vis (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International journal of remote sensing*, 16(12):2165–2185, 1995.

M. Ringnér. What is principal component analysis? *Nature Biotechnology*, 26(3):303–304, Mar. 2008. ISSN 1087-0156, 1546-1696.

S. Roessner, K. Segl, U. Heiden, and H. Kaufmann. Automated differentiation of urban surfaces based on airborne hyperspectral imagery. *IEEE Transactions on Geoscience and Remote sensing*, 39(7):1525–1532, 2001.

L. C. Rowan, A. F. Goetz, and R. P. Ashley. Discrimination of hydrothermally altered and unaltered rocks in visible and near infrared multispectral images. *Geophysics*, 42(3): 522–535, 1977.

L. Y. Sahib. *Crude oil and oil brine seeps: sources, detection and environmental effects in soil and water, Kirkuk NE Iraq*. PhD thesis, Technische Universität, Darmstadt, Januar 2019.

C. Schinke, P. Christian Peest, J. Schmidt, R. Brendel, K. Bothe, M. R. Vogt, I. Kröger, S. Winter, A. Schirmacher, S. Lim, et al. Uncertainty analysis for the coefficient of band-to-band absorption of crystalline silicon. *AIP Advances*, 5(6):067168, 2015.

S. Schmidtlein and J. Sassin. Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. *Remote Sensing of Environment*, 92(1):126–138, July 2004. ISSN 0034-4257.

S. Schmidtlein, P. Zimmermann, R. Schüpferling, and C. Weiss. Mapping the floristic continuum: Ordination space position estimated from imaging spectroscopy. *Journal of Vegetation Science*, 18(1):131–140, 2007.

S. Schmidtlein, C. Oldenburg, H. Feilhauer, and B.-H. Mevik. Package 'autopls'. *Partial least square regression with backward selection of predictors.[WWW Document]. URL https://cran. r-project. org/web/packages/autopls/index. html (accessed 7.1. 17)*, 2015.

J. Schopfer, S. Dangel, M. Kneubühler, and K. I. Itten. The improved dual-view field goniometer system figos. *Sensors*, 8(8):5120–5140, 2008.

C. Schuster, J. Honold, S. Lauf, and T. Lakes. Urban heat stress: Novel survey suggests health and fitness as future avenue for research and adaptation strategies. *Environmental Research Letters*, 12(4):044021, 2017.

P. Schwind, M. Schneider, and R. Müller. Improving hyspex sensor co-registration accuracy using brisk and sensor-model based ransac. *ISPRS Archives*, 40:371–376, 2014.

K. Segl, S. Roessner, U. Heiden, and H. Kaufmann. Fusion of spectral and shape features for identification of urban surface cover types using reflective and thermal hyperspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2):99–112, 2003.

K. Segl, L. Guanter, C. Rogass, T. Kuester, S. Roessner, H. Kaufmann, B. Sang, V. Mogulsky, and S. Hofer. EeteS—The EnMAP End-to-End Simulation Tool. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(2):522–530, Apr. 2012. ISSN 1939-1404.

K. C. Seto and J. M. Shepherd. Global urban land-use trends and climate impacts. *Current Opinion in Environmental Sustainability*, 1(1):89–95, 2009.

K. C. Seto, R. Sánchez-Rodríguez, and M. Fragkias. The new geography of contemporary urbanization and the environment. *Annual review of environment and resources*, 35: 167–194, 2010.

Z. Shao, H. Fu, D. Li, O. Altan, and T. Cheng. Remote sensing monitoring of multi-scale watersheds impermeability for urban hydrological evaluation. *Remote Sensing of Environment*, 232:111338, 2019.

S. Silvestre, A. Tahri, F. Tahri, S. Benlebna, and A. Chouder. Evaluation of the performance and degradation of crystalline silicon-based photovoltaic modules in the saharan environment. *Energy*, 152:57–63, 2018.

S. Skowronek, R. Van De Kerchove, B. Rombouts, R. Aerts, M. Ewald, J. Warrie, F. Schiefer, C. Garzon-Lopez, T. Hattab, O. Honnay, J. Lenoir, D. Rocchini, S. Schmidtlein, B. Somers, and H. Feilhauer. Transferability of species distribution models for the detection of an invasive alien bryophyte using imaging spectroscopy data. *International Journal of Applied Earth Observation and Geoinformation*, 68:61–72, June 2018. ISSN 0303-2434.

C. Small. A global analysis of urban reflectance. *International Journal of Remote Sensing*, 26(4):661–681, 2005.

C. Small, A. Okujeni, S. v. d. Linden, and B. Waske. 6.07 - Remote Sensing of Urban Environments. In S. Liang, editor, *Comprehensive Remote Sensing*, pages 96–127. Elsevier, Oxford, 2018. ISBN 978-0-12-803221-3. doi: https://doi.org/10.1016/B978-0-12-409548-9.10380-X. URL https://www.sciencedirect.com/science/article/pii/B978012409548910380X.

B. Somers, G. P. Asner, L. Tits, and P. Coppin. Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment*, 115(7):1603–1616, 2011.

S. Sorrell. Reducing energy demand: A review of issues, challenges and approaches. *Renewable and Sustainable Energy Reviews*, 47:74–82, 2015.

H. Su, W. Yao, Z. Wu, P. Zheng, and Q. Du. Kernel low-rank representation with elastic net for china coastal wetland land cover classification using gf-5 hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171:238–252, 2021.

J. Tao and S. Yu. Review on feasible recycling pathways and technologies of solar photovoltaic modules. *Solar Energy Materials and Solar Cells*, 141:108–124, 2015.

J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science (New York, N.Y.)*, 290(5500):2319–2323, Dec. 2000. ISSN 0036-8075.

W. R. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46(sup1):234–240, June 1970. ISSN 0013-0095.

R. Upreti, Z.-H. Wang, and J. Yang. Radiative shading effect of urban trees on cooling the regional built environment. *Urban Forestry & Urban Greening*, 26:18–24, 2017.

S. Van der Linden, A. Rabe, M. Held, B. Jakimow, P. J. Leitão, A. Okujeni, M. Schwieder, S. Suess, and P. Hostert. The enmap-box—a toolbox and application programming interface for enmap data processing. *Remote Sensing*, 7(9):11249–11266, 2015.

S. van der Linden, A. Okujeni, F. Canters, J. Degerickx, U. Heiden, P. Hostert, F. Priem, B. Somers, and F. Thiel. Imaging spectroscopy of urban environments. *Surveys in Geophysics*, 40(3):471–488, May 2019. ISSN 1573-0956.

I. Visa, B. Burduhos, M. Neagoe, M. Moldovan, and A. Duta. Comparative analysis of the infield response of five types of photovoltaic modules. *Renewable Energy*, 95:178–190, 2016.

J.-F. Wang, A. Stein, B.-B. Gao, and Y. Ge. A review of spatial sampling. *Spatial Statistics*, 2:1–14, 2012.

K. J. Ward, S. Chabrillat, M. Brell, F. Castaldi, D. Spengler, and S. Foerster. Mapping soil organic carbon for airborne and simulated enmap imagery using the lucas soil database and a local plsr. *Remote Sensing*, 12(20):3451, 2020.

Y. Wen, G. Schoups, and N. Van De Giesen. Organic pollution of rivers: Combined threats of urbanization, livestock farming and global climate change. *Scientific reports*, 7(1):1–9, 2017.

Q. Weng. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, 117:34–49, 2012.

Q. Weng and D. A. Quattrochi. *Urban remote sensing*. CRC press, 2018.

E. A. Wentz, S. Anderson, M. Fragkias, M. Netzband, V. Mesev, S. W. Myint, D. Quattrochi, A. Rahman, and K. C. Seto. Supporting global environmental change research: A review of trends and knowledge gaps in urban remote sensing. *Remote Sensing*, 6(5):3879–3905, 2014.

R. Wittig, H. Sukopp, and B. Klausnitzer. *Die ökologische Gliederung der Stadt*. Gustav Fischer Verlag, July 1998. ISBN 978-0-444-53869-7.

J. R. Wolch, J. Byrne, and J. P. Newell. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and urban planning*, 125:234–244, 2014.

C. Wolfram, O. Shelef, and P. Gertler. How will energy demand develop in the developing world? *Journal of Economic Perspectives*, 26(1):119–38, 2012.

Q. Xie, M. Zhou, Q. Zhao, D. Meng, W. Zuo, and Z. Xu. Multispectral and hyperspectral image fusion by ms/hs fusion net. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1585–1594, 2019.

H. Yan, X. Wang, P. Hao, and L. Dong. Study on the microclimatic characteristics and human comfort of park plant communities in summer. *Procedia Environmental Sciences*, 13:755–765, 2012.

N. Yokoya, C. Grohnfeldt, and J. Chanussot. Hyperspectral and multispectral data fusion: A comparative review of the recent literature. *IEEE Geoscience and Remote Sensing Magazine*, 5(2):29–56, 2017.

J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal. DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States. *Joule*, 2(12): 2605–2617, 2018.

J. Yuan, H.-H. L. Yang, O. A. Omitaomu, and B. L. Bhaduri. Large-scale solar panel mapping from aerial images using deep convolutional networks. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 2703–2708. IEEE, 2016.

A. Zare and K. Ho. Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing. *IEEE Signal Processing Magazine*, 31(1):95–104, 2013.

Y. Zha, J. Gao, and S. Ni. Use of normalized difference built-up index in automatically mapping urban areas from tm imagery. *International journal of remote sensing*, 24(3): 583–594, 2003.

D. Zhang, G. Liu, C. Chen, Y. Zhang, Y. Hao, and M. Casazza. Medium-to-long-term coupled strategies for energy efficiency and greenhouse gas emissions reduction in beijing (china). *Energy Policy*, 127:350–360, 2019.

D.-L. Zhang, Y.-X. Shou, R. R. Dickerson, and F. Chen. Impact of upstream urbanization on the urban heat island effects along the washington–baltimore corridor. *Journal of applied meteorology and climatology*, 50(10):2012–2029, 2011.

J. Zhang, B. Rivard, A. Sánchez-Azofeifa, and K. Castro-Esau. Intra-and inter-class spectral variability of tropical tree species at la selva, costa rica: Implications for species identification using hydice imagery. *Remote Sensing of Environment*, 105(2):129–141, 2006.

L. Zhu, J. Suomalainen, J. Liu, J. Hyyppä, H. Kaartinen, H. Haggren, et al. A review: Remote sensing sensors. *Multi-purposeful application of geospatial data*, pages 19–42, 2018.

Z. Zhu. Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130:370–384, 2017.