Remote Sensing of Environment 321 (2025) 114657

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A gradient-based 3D nonlinear spectral model for providing components optical properties of mixed pixels in shortwave urban images

Zhijun Zhen [a,b], Shengbo Chen [a,c,*], Nicolas Lauret [b], Abdelaziz Kallel [d], Eric Chavanon [b], Tiangang Yin [e], Jonathan León-Tavares [f], Biao Cao [g], Jordan Guilleux [b], Jean-Philippe Gastellu-Etchegorry [b,*]

[a] *The College of Geoexploration Science and Technology, Jilin University, Changchun 130026, China*
[b] *CESBIO, Université de Toulouse, UT3-Paul Sabatier/CNES/CNRS/INRAE/IRD, 18, Avenue Edouard Belin, 31401 Toulouse, France*
[c] *Jilin Institute of GF Remote Sensing Application, Changchun 130012, China*
[d] *Digital Research Center of Sfax, Technopole of Sfax, Sakiet Ezzit 3021, Tunisia*
[e] *Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong*
[f] *Vlaamse Instelling voor Technologisch Onderzoek (VITO), Boeretang 200, 2400 Mol, Belgium*
[g] *Innovation Research Center of Satellite Application, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China*

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
DART calibration
Monospectral image
Multispectral image
Spectral unmixing
Urban meteorology
Vegetation

ABSTRACT

Unmixing optical properties (OP) of land covers from coarse spatial resolution images is crucial for microclimate and energy balance studies. We propose the Unmixing Spectral method using Discrete Anisotropic Radiative Transfer (DART) model (US-DART), a novel approach for unmixing endmember OP in the shortwave domain from mono- or multispectral remotely sensed images. US-DART comprises four modules: pure pixel selection, linear spectral mixture analysis, gradient iterations, and spectral correlation. US-DART requires a surface reflectance image, a 3D mock-up with facets' group information, and standard DART parameters (e.g., spatial resolution and skylight ratio) as inputs, producing an OP map for each scene element. The accuracy of US-DART is evaluated using two types of scenes (vegetation and urban) and images (Sentinel-2 surface reflectance and DART-simulated pseudo-satellite images). Results demonstrate a median relative error of approximately 0.1 % for pixel reflectance, with higher accuracy for opaque surfaces compared to translucent materials. Excluding co-registration errors and sensor noise, the median relative error of OP is typically around 1 % for opaque elements and 1–5 % for translucent elements with an accurate a priori "reflectance-transmittance" ratio. US-DART enhances our ability to derive detailed OP from coarse-resolution imagery, potentially enabling more accurate modeling of spatial resolution conversions, and energy dynamics, including albedo and shortwave radiation balance, across diverse environments.

1. Introduction

The optical properties (OP) of land covers are crucial in microclimate and energy balance studies (Chen and Liu, 2020; Shen et al., 2023), in monitoring vegetation biochemistry (Chen et al., 2022; Liu et al., 2023; Xu et al., 2023), and in simulation of time-series data for albedo and radiation balance (Leng et al., 2023; Wang et al., 2022). Due to its high frequency and global observations, remote sensing is extensively utilized for retrieving OP (Wang et al., 2024a); however, the pixel dimensions of satellite images often exceed those of land cover elements, as capturing urban structures accurately requires a spatial resolution of 0.5 to 10 m (Welch, 1982), which is finer than the current spatial resolutions, so in widely used satellite images such as those from Sentinel-2 (10 m) and Landsat 8 (30 m) (Dong et al., 2023), the pixel signal typically results from light scattering by multiple scene elements (Keshava and Mustard, 2002), implying that pixels, irrespective of their size, cannot be entirely homogeneous (Xie et al., 2021; Xie et al., 2022). This heterogeneity, or mixed pixels, coupled with multiple scattering, complicates the retrieval of land cover OP from remote sensing images.

Spectral mixture analysis serves as an effective method to tackle the issue of mixed pixels (Guo et al., 2024) by employing either a linear or a nonlinear model to convert the pixel value into a set of endmembers and

---

* Corresponding authors at: The College of Geoexploration Science and Technology, Jilin University, Changchun 130026, China (S. Chen).
E-mail addresses: chensb@jlu.edu.cn (S. Chen), jean-philippe.gastellu@iut-tlse3.fr (J.-P. Gastellu-Etchegorry).

https://doi.org/10.1016/j.rse.2025.114657
Received 21 September 2024; Received in revised form 3 February 2025; Accepted 14 February 2025

their corresponding abundance. Linear spectral unmixing proves efficient in simple environments with minimal multi-scattering, while nonlinear spectral unmixing enhances the precision of component identification and abundance estimation in heterogeneous scenes, such as cities, where multi-scattering is substantial (Zhen et al., 2021; Zhu and Shi, 2018; Zhu et al., 2012). The spectral variability of the endmember, also known as the multi-endmembers problem, poses a significant challenge for unmixing in urban areas from spectral confusion (e.g., intra-class and inter-class variability), illumination conditions (e.g., topography), and atmospheric effects (e.g., the spatial variation of the atmospheric properties in the image) (Borsoi et al., 2021; Drumetz et al., 2019a). This paper focuses on spectral confusion and illumination conditions.

Spectral confusion, a consequence of the spatial variability of endmembers, is divided into intra-class and inter-class variability (Zhang et al., 2006), and unmixing methods that take these variabilities into account fall into four categories: endmember bundles, local spectral unmixing, computational models, and parametric physics-based models (Drumetz et al., 2019b).

Endmember bundles (Roberts et al., 1998) enhance diversity by substituting each endmember's signature with multiple candidates, though their efficiency depends on the precision of bundle extraction, which can be challenging. For example, the study (Somers et al., 2012) addresses the inherent limitations of conventional endmember extraction algorithms (EEAs) in hyperspectral image analysis by proposing a novel methodology for automated spectral bundle generation that accounts for endmember variability through iterative sampling and clustering techniques, thereby enhancing the accuracy of fractional abundance estimations in spectral unmixing applications. Local spectral unmixing (Achanta et al., 2012) adaptively segments image regions to mitigate endmember variability, but integrating local results into a global context is difficult.

Computational models (Du et al., 2014) allow endmembers to vary locally by assigning statistical distributions, yet they fail to account for the physical processes causing intraclass variability. For example, the study (Thouvenin et al., 2015) propose a novel linear mixing model and corresponding alternating direction method of multipliers-based optimization algorithm for hyperspectral unmixing that explicitly incorporates both spatial and spectral endmember variabilities, thereby enabling more accurate estimation of reference spectral signatures, their abundance fractions, and inherent variability patterns within remotely sensed imagery. The study (Brezini et al., 2021) presents an enhanced Nonnegative Matrix Factorization-based spectral unmixing methodology that addresses intra-class variability in hyperspectral remote sensing through the incorporation of two novel penalization terms into a structured additively-tuned linear mixing model's cost function, yielding improved management of endmember spectral variations via iterative multiplicative update rules.

Physics-based models (Hapke, 2012) utilize a simplified radiative transfer model to interpret variations with physically interpretable parameters, but they are complex, pose inversion challenges, and depend on numerous hard-to-determine parameters. For example, to address the challenge of spectral variability in hyperspectral imagery, the study (Hong et al., 2018) introduces an augmented Linear Mixing Model (LMM) that incorporates a data-driven learning strategy and spectral variability dictionary to improve the accuracy of abundance map estimation in spectral unmixing problems. The study (Imbiriba et al., 2018) introduces a generalized linear mixing model (GLMM) that extends the capabilities of the extended linear mixing model (ELMM) to better handle complex spectral distortions in hyperspectral image analysis by allowing for uneven effects across different wavelength intervals.

Illumination conditions (e.g., shadows) influence reflectance magnitude and cause spectral distortion, leading to endmember spectral variability (Uezato et al., 2020), as sunlit pixels receive both direct and diffuse solar radiation, whereas shaded pixels only receive diffuse radiation (Zhang et al., 2023). Current research addresses this spectral variability by incorporating shadow endmembers or considering illumination conditions (Uezato et al., 2020). One approach introduces either a single (Nascimento and Dias, 2005) or multiple (Uezato et al., 2015) shadow endmembers, or adds extra variation terms in the mixing model to simulate lighting changes (Drumetz et al., 2016). However, these models oversimplify the mechanisms of spectral variation and fail to adequately describe light interactions within pixels and the causes of spectral changes, leading to potential misinterpretation of spectral variations caused by illumination conditions or material physical properties (Uezato et al., 2016).

Models that account for illumination conditions consider the effects of both direct sunlight and diffuse skylight by integrating Light Detection and Ranging (LiDAR)-derived digital surface model (DSM) data (Uezato et al., 2016). The DSM, for instance, is adaptable to varying illumination conditions and can be effectively used for spectral mixing analysis. When combined with internal data, DSM allows for the computation of spatial regularization weights, which helps identify and preserve edge structures in abundance maps, thereby enhancing unmixing quality (Uezato et al., 2018). Moreover, radiance and DSM data can be used together to unmix hyperspectral images, taking into account the effects of variable illumination and shadows (Uezato et al., 2020).

Machine learning offers new avenues for addressing the multi-endmember problem of spectral unmixing in complex urban areas. Unrolled optimization-based neural networks (Zhou and Rodrigues, 2021) prove advantageous for unmixing, particularly when considering nonlinearity (Wang et al., 2019) and endmember variability (Borsoi et al., 2019). Physically motivated machine learning approaches have been successfully applied to unmixing (Qian et al., 2020), providing interpretability for the estimated endmembers and abundance parameters. Due to the intrinsic low-dimensionality of the abundance space (Palsson et al., 2018), several autoencoder approaches have been proposed to tackle phenomena such as nonlinearity (Li et al., 2021), endmember variability (Shi et al., 2021), and outliers (Su et al., 2019). However, these methods fall short of accounting for existing spectral variability together with nonlinear effects (Borsoi et al., 2023).

Nonlinear unmixing models primarily address microscopic nonlinear phenomena, neglecting macroscopic multiple scattering between geometrically diverse structures—particularly in vertically heterogeneous urban environments. The core limitation of abundance-based nonlinear mixing models lies in the nonlinear relationship between physical properties and measured reflectance. This manifests through: (1) reflectance-physical property misalignment, where reflectance correlates nonlinearly with material mass and cross-sectional area; and (2) the "small but bright" effect, wherein minor highly-reflective objects disproportionately influence spectral signatures compared to larger, less-reflective objects within a pixel. Consequently, while area proportionsparadoxically retain accuracy in assessing materials' contributions to aggregate reflectivity measurements, they inadequately characterize reflectance contributions (Bioucas-Dias et al., 2012).

Mitsuba 3 (Vicini et al., 2021) introduces the three-dimensional radiative transfer equation into the design of deep learning loss functions, establishing a physics-constrained inversion framework without training samples. It transfers gradients among parameters with clear physical significance. However, although Mitsuba 3 is widely used in computer graphics, its adaptation to remote sensing applications has been limited due to the complexity of remote sensing imaging conditions compared with computer graphics. For instance, satellite images can be significantly inaccurate due to various noise sources, such as sensors and the atmosphere. Additionally, the 3D mock-ups used in remote sensing models often deviate significantly from real-world conditions due to the necessary simplifications applied to remote sensing scenes. These simplifications can result in inaccuracies stemming from imprecise measurements, which, in turn, adversely affect the performance of the inversion algorithm. Additionally, the inefficiency of the gradient descent method and its reliance on numerous iterations make it unsuitable for the modern demands of processing large-scale remote sensing data and conducting extensive scenario simulations.

The Discrete Anisotropic Radiative Transfer (DART) model calibration (Landier et al., 2018) exemplifies the use of a 3D radiative transfer model for the inversion of urban element OP from the remote sensing images. It employs an urban 3D mock-up containing endmember shape and classification information for nonlinear unmixing, enabling the estimation of endmember OPs while distinguishing spectral variability caused by illumination conditions or spectral confusion. This approach can generate an OP map per endmember at a resolution comparable to satellite imagery, and these OP maps have been successfully utilized for urban energy modeling (Chrysoulakis et al., 2018). Although the model effectively addressed the multiple scattering of the three-dimensional structure when extracting pure pixels' OP, achieving good results. However, the method is less effective in considering mixed pixels. Existing studies have shown that the spatial consistency of these OP maps is subpar, particularly in coarse mixed pixels with complex composition (Zhen et al., 2021).

Spectral gradients may be a potential replacement of abundance for unmixing. They establish direct correlations with the fundamental physics of light-material interactions, more accurately capture the inherent nonlinearity of spectral mixing processes, and circumvent the problematic assumption of linear relationships between abundance and reflectance measurements. This theoretical framework provides a more rigorous foundation for analyzing complex spectral mixing phenomena in remote sensing applications. Therefore, to enhance the accuracy of DART calibration on mixed pixels, we propose the Unmixing Spectral method using DART (US-DART) by incorporating spectral gradients. This method amalgamates knowledge from computer graphics, deep learning, and geostatistics to augment the retrieval of endmember OP. As a physical model, US-DART transfers gradients among parameters with explicit physical significance. We evaluated the accuracy of US-DART using vegetation and urban mock-ups, along with Sentinel-2 reflectance images and DART-simulated pseudo satellite images. The contributions of US-DART can be encapsulated compared with DART calibration as follows:

- A new Pure Pixel Selection (PPS) module enhances the initial OP value for equation system solving in the Linear Spectral Mixture Analysis (LSMA) module by utilizing information from the satellite image.
- A new Gradient Iteration (GI) module replaces the previous Newton-bisection method by improving the treatment of multiple scattering. In contrast to the Newton-bisection method, which operates on a pixel-by-pixel basis, the GI module uses multi-pixel methods. This approach allows for a more comprehensive consideration of the spatial consistency of endmember OP.
- A new Spectral Correlation (SC) module leverages OP features and optional prior knowledge to eliminate erroneous pixels in the OP maps.
- The process of equation system solving is enhanced by incorporating judgments of stability conditions, such as determinant, condition number, and spectral radius of the equation system, to limit the generation of erroneous pixels in the OP maps.
- Multiple threading is implemented on bands for multispectral images or image segments for mono-spectral images.

2. Methodology

Section 2.1 introduces the essential background knowledge pertaining to DART and its calibration process. Section 2.2 elucidates the framework and intricate modules of the US-DART method. Section 2.3 presents the data used and the preprocessing steps undertaken for the accuracy assessment of the US-DART method.

2.1. Theory background

Section 2.1.1 provides the background information on DART. Section 2.1.2 explores the necessary details regarding the DART calibration.

2.1.1. DART model

The DART model (Gastellu-Etchegorry, 2008) is a comprehensive and precise 3D radiative transfer model based on physics, designed to track the propagation of shortwave (SW) and longwave (LW) radiation within the Earth-atmosphere system. It provides two methods for simulating radiative transfer. The first method, known as the discrete ordinate method (Yin et al., 2013) or DART-Flux Tracing (Gastellu-Etchegorry et al., 2015), traces radiation in a finite number of directions within a voxelized landscape. The second method, a bi-directional Monte Carlo algorithm called DART-Lux (Wang et al., 2024b), tracks photons in any direction within an unvoxelized landscape, thereby improving simulation efficiency in terms of computation time and memory usage. DART is freely accessible at the University of Toulouse (dart.omp.eu) for non-commercial academic and research use.

DART can simulate the 3D radiative budget as well as the SW and LW observations—such as reflectance and brightness temperature—of imaging radiometers and LiDAR scanners on space, airborne, and in-situ platforms. This capability extends to various urban and natural landscapes and accommodates any sun/sensor/atmosphere configuration, including different spatial and spectral resolutions, sensor viewing directions, and platform altitudes.

DART is capable of both simulating and importing landscapes. These landscapes can range from realistic representations, such as 3D urban geometric databases, to more schematic depictions, like the superimposition of homogeneous and horizontal layers of a turbid medium (i.e., a random distribution of infinitely small facets). The elements of these scenes can be simulated as either Lambertian or non-Lambertian facets or as volumes of a turbid medium. Furthermore, DART is capable of simulating the atmosphere for any gas and aerosol content, taking into account various spectral properties such as phase functions, vertical profiles, extinction coefficients, and spherical albedo.

The DART model was employed in two distinct ways. First, it was used to create a homogeneous vegetation canopy and simulate pseudo-satellite images, providing data for the accuracy assessment of TRUST-DART. Second, it computed the gradient for each element within a pixel. DART generated spectral element radiance images for each type of element, offering insights into the contribution of each element within each pixel. Consequently, DART proved invaluable for mixing and unmixing processes.

2.1.2. DART calibration

DART calibration (Landier, 2018; Landier et al., 2018) is a spectrum unmixing technique designed to extract the per-pixel OP of each urban element from satellite images within the SW domain, utilizing the DART model. The process employs an iterative methodology that continuously refines the input OP by comparing the DART model outputs with actual satellite images. This process is divided into two main steps:

The procedure begins with LSMA, employing multi-pixel methodologies to characterize spectral components. The inversion process is initialized using user-defined OP values for each element classification, potentially sourced from the DART optical property database. An initial OP distribution map is generated under linear unmixing assumptions, operating on the principle that identical elements maintain consistent OP within a defined sampling window.

Subsequently, the dichotomous-Newton method is employed to refine the OP derived from the LSMA module by incorporating the multi-scattering effect. This process is iterative and conducted pixel-by-pixel without considering the interdependence between spectral bands. If the pixel reflectance of the DART simulated images is lower than that of the satellite image, the OP values of the elements in this pixel are increased, and vice versa. The procedure concludes when the DART simulated image, with the OP map obtained through inversion, closely matches the satellite image.

2.2. Unmixing spectral method using the DART model (US-DART)

Section 2.2.1 provided an overview of the US-DART method. Subsequently, each module of US-DART was detailed in Sections 2.2.2 through 2.2.5.

2.2.1. Overview

The US-DART method was designed to extract the OP map of each endmember from mono or multispectral surface reflectance images. The input parameters were a surface reflectance image, a 3D mock-up with element classification information, and common DART parameters (e.g., spatial resolutions and skylight ratio). The output generated was an OP map for each element. The algorithm comprised four modules: PPS, LSMA, GI, and SC (Fig. 1). The PPS module extracted the initial OP of element i at wavelength λ, denoted as $\rho_{i,\lambda}^{\mathrm{PPS}}$, from the pure pixels in the satellite image. This initial OP was then refined to $\rho_{i,\lambda}^{\mathrm{LSMA}}$ using the LSMA module. Finally, considering multiple scattering, $\rho_{i,\lambda}^{\mathrm{LSMA}}$ was iteratively refined to $\rho_{i,\lambda}^{\mathrm{GI}}$ using the GI module until the DART simulated image was nearly the same as the satellite image. Throughout this process, the SC module was involved in all other modules to detect and correct any erroneous pixels. The subsequent subsections provided detailed descriptions of these four modules.

2.2.2. Pure pixel selection module

The objective of the PPS module was to derive OP $\rho_{i,\lambda}^{\mathrm{PPS}}$ for each element i from satellite image $R_{\mathrm{sat},\lambda}$ at wavelength λ. Initially, a DART simulation was executed to compute the fraction of element i in every pixel. Subsequently, pure pixels for each i were selected based on a defined fraction threshold of 0.9. If no pure pixels were able to be identified under this threshold, this threshold was gradually reduced from 0.9 to 0.1 in reductions of 0.1. If no pure pixels were found even at the minimum threshold (0.1), $\rho_{i,\lambda}^{\mathrm{PPS}}$ was assigned the user-input OP. Following this, $\rho_{i,\lambda}^{\mathrm{PPS}}$ of mixed pixels were assigned using the $\rho_{i,\lambda}^{\mathrm{PPS}}$ from their nearest pure pixels. If multiple pure pixels were found at an equal distance, the median value of all such pixels was utilized for the assignment. Due to co-registration errors between DART and satellite images, some pixels at the boundary of pure pixel zones possibly contained signals from other elements. Consequently, an optional erosion process was able to be employed for high-resolution images with a significant number of pure pixels to purify the selected pixels further if needed (Cubero-Castan et al., 2014).

The term OP denoted the reflectance for opaque and the albedo for translucent elements, such as leaves, which was the sum of reflectance and transmittance. The OP approximately equaled the pixel reflectance for pure pixels composed of opaque horizontal elements without multiple scattering. For translucent materials, their pixel reflectance depended on the OP of the elements and the effects of multiple scattering beneath them. As a rough estimation, we employed Eq. (1) to derive OP of translucent materials, which underwent refinement in a subsequent module:

$$\rho_{i,\lambda}^{\mathrm{PPS}} = \sqrt{R_{\mathrm{sat},\lambda}} \tag{1}$$

In our practice, this estimation was more effective than canopy reflectance directly as leaf albedo. To distinguish between reflectance and transmittance from albedo for leaves, we assumed that the ratio of reflectance to transmittance was known beforehand.

2.2.3. Linear spectral mixture analysis module

The LSMA module was designed to refine OP from $\rho_{i,\lambda}^{\mathrm{PPS}}$ to $\rho_{i,\lambda}^{\mathrm{LSMA}}$ using LSMA, regardless of multiple scattering between elements. We want DART simulated image, which is the sum of all reflectance images of each element, very close to the $R_{\mathrm{sat},\lambda}$. Consequently, $R_{\mathrm{sat},\lambda}$ was expressed as:

$$R_{\mathrm{sat},\lambda} = \sum_{i=1}^{I} R_{i,\lambda} = \sum_{i=1}^{I} \left(G_{i,\lambda}^{\mathrm{LSMA}} \cdot \rho_{i,\lambda}^{\mathrm{LSMA}} \right) \tag{2}$$

where I represented the total number of elements, $R_{i,\lambda}$ was the reflectance image of element i in wavelength λ, and $G_{i,\lambda}^{\mathrm{LSMA}}$ was the gradient calculated by the DART model with $\rho_{i,\lambda}^{\mathrm{PPS}}$, which was determined by calculating the differential pixel reflectance for the OP of the elements.

In the presence of mixed pixels (i.e., several unknowns and a unique equation), the solution adopted was to increase the number of equations by incorporating nearby pixels. This incorporation was achieved by considering a moving rectangular window of M pixels, under the assumption that the same elements in the sampling window shared the

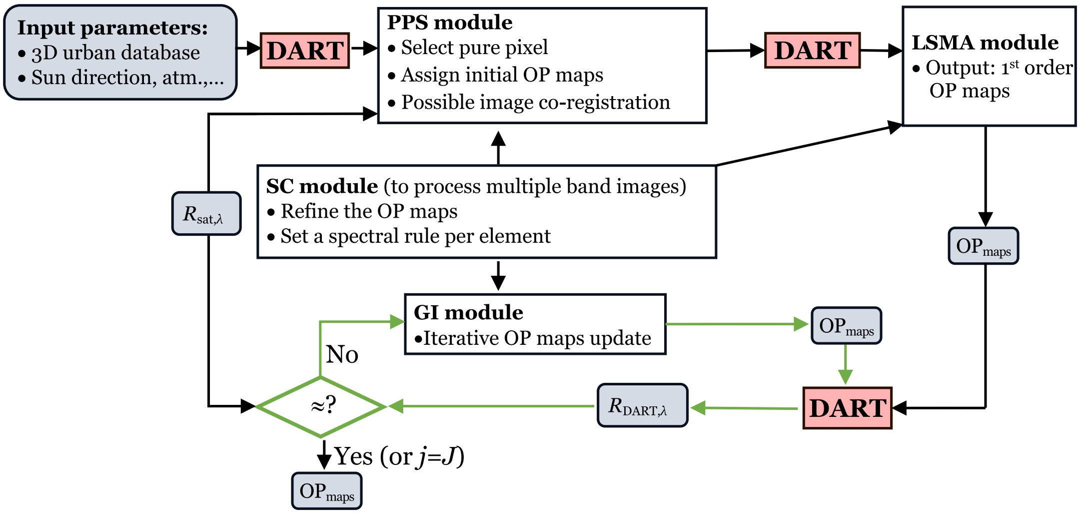


Fig. 1. Algorithm flow chart of US-DART: input and output parameters (rounded rectangles), processes (square corner rectangles), decision (rhombus), DART simulation (light red), and iterative process (green). $R_{\mathrm{sat},\lambda}$: satellite reflectance image at wavelength λ; (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
$R_{\mathrm{DART},\lambda}$: DART reflectance image at wavelength λ; $\mathrm{OP}_{\mathrm{maps}}$: optical properties maps; PPS: pure pixel selection module; LSMA: linear spectral mixing analysis module; GI: gradient iteration module; SC: spectral correlation module.

same OP based on spatial continuity. This extension of Eq.(2) to all the pixels of the window resulted in a total of I elements and M equations. To ensure the equation system Eq.(3) was not underdetermined, the number of pixels M in the window needed to be larger than or equal to I. Users had the flexibility to customize the size and shape of each sampling window.

$$\begin{pmatrix} R_{\text{sat},\lambda}(1) \\ \cdots \\ R_{\text{sat},\lambda}(M) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{I}\left[G_{i,\lambda}^{\text{LSMA}}(1)\cdot\rho_{i,\lambda}^{\text{LSMA}}(1)\right] \\ \cdots \\ \sum_{i=1}^{I}\left[G_{i,\lambda}^{\text{LSMA}}(M)\cdot\rho_{i,\lambda}^{\text{LSMA}}(M)\right] \end{pmatrix} \tag{3}$$

Depending on the distribution pattern of the elements, the spatial resolution of the sensor, and the shape of the sampling window, the equation system formed by different sampling windows exhibited distinct stability, typically indicated using mathematical terms such as condition number (set here as 50,000) (Cline et al., 1979) and spectral radius (set here as 1.1) (Dunford et al., 1963). Thus, users were able to define associated thresholds to decide whether to solve or skip an equation system based on its stability. In addition, when solving the equation system, the value of $\rho_{i,\lambda}^{\text{PPS}}$ was used as the initial value, and a constraint condition was applied to limit the solved $\rho_{i,\lambda}^{\text{LSMA}}$ within a range. The condition was set as $(1 \pm \varepsilon) \times \rho_{i,\lambda}^{\text{PPS}}$ (ε set here as 0.5). Following the LSMA module, a median filter was employed to fill the $\rho_{i,\lambda}^{\text{LSMA}}$ for the pixels that were not solved.

2.2.4. Gradient iteration module

The objective of the GI module was to further refine OP from $\rho_{i,\lambda}^{\text{LSMA}}$ to $\rho_{i,\lambda}^{\text{GI}}$ by taking into account multi-scattering. Based on the definition of the gradient, the reflectance image simulated by DART at the next iteration ($j+1$), denoted as $R_{\text{DART},\lambda}^{j+1}$, was given by:

$$R_{\text{DART},\lambda}^{j+1} = R_{\text{DART},\lambda}^{j} + \sum_{i=1}^{I}\left[G_{i,\lambda}^{j}\cdot\left(\rho_{i,\lambda}^{j+1} - \rho_{i,\lambda}^{j}\right)\right] \tag{4}$$

Here, $R_{\text{DART},\lambda}^{j}$ was the DART simulated image at the current iteration j, and $G_{i,\lambda}^{j}$ was the gradient calculated in the GI module. The goal was to make the simulated reflectance image nearly identical to the satellite image, i.e., $R_{\text{DART},\lambda}^{j+1} = R_{\text{sat},\lambda}$. Therefore, Eq.(4) was rewritten as:

$$R_{\text{sat},\lambda} = R_{\text{DART},\lambda}^{j} + \sum_{i=1}^{I}\left[G_{i,\lambda}^{j}\cdot\left(\rho_{i,\lambda}^{j+1} - \rho_{i,\lambda}^{j}\right)\right] \tag{5}$$

Extending Eq. (5) to the M pixels of the moving window led to:

$$\begin{pmatrix} R_{\text{sat},\lambda}(1) \\ \cdots \\ R_{\text{sat},\lambda}(M) \end{pmatrix} = \begin{pmatrix} R_{\text{DART},\lambda}^{j}(1) \\ \cdots \\ R_{\text{DART},\lambda}^{j}(M) \end{pmatrix} - \begin{pmatrix} \sum_{i}^{I} G_{i,\lambda}^{j}(1)\cdot\rho_{i}^{j}(1) \\ \cdots \\ \sum_{i}^{I} G_{i,\lambda}^{j}(M)\cdot\rho_{i}^{j}(M) \end{pmatrix} + \begin{pmatrix} G_{1,\lambda}^{j}(1) & \cdots & G_{I,\lambda}^{j}(1) \\ \cdots & \cdots & \cdots \\ G_{1,\lambda}^{j}(M) & \cdots & G_{I,\lambda}^{j}(M) \end{pmatrix} \cdot \begin{pmatrix} \rho_{1}^{j+1} \\ \cdots \\ \rho_{I}^{j+1} \end{pmatrix} \tag{6}$$

Finally, solving Eq.(6) yielded the unknowns $\rho_{i,\lambda}^{j+1}$ for the I elements. The method to solve this equation system is similar to Eq. (3).

2.2.5. Spectral correlation module

The SC module was designed to enhance the reliability of the output of PPS, LSMA, and GI modules by identifying and rectifying inaccurate pixels in the OP map. This process utilized the reference spectral signatures of each endmember. For instance, the reflectance of green vegetation was expected to be higher in the near-infrared (NIR) band compared to the red band. Users had the flexibility to customize spectral rules for each element. These rules accommodated operators on spectral bands, with priority determined by parentheses. The operators included arithmetic ones: addition (+), subtraction (−), multiplication (*), and division (/); relational ones: greater than (>), greater than or equal to (≥), less than (<), and less than or equal to (≤); and logical ones: and (&), or (|). Pixels whose OP were inconsistent with the spectral rules were assigned values using a median filter.

2.3. Data

We evaluated the accuracy of US-DART using Sentinel-2 and DART-simulated pseudo satellite images, focusing on vegetation and urban scenes. These scenes were detailed in Section 2.3.1. The OP of the scene elements, which were utilized to simulate the pseudo satellite images, were presented in Section 2.3.2. Finally, Section 2.3.3 outlined the satellite data and their preprocessing procedures.

2.3.1. The 3D vegetation and urban scenes

(1) **Schematic vegetation scenes**.

DART created the vegetation scene as three volumes filled with a turbid medium (representing vegetation plots) and three bare ground plots (Fig. 2). The image simulated by DART served as a pseudo satellite image, which was used to assess the accuracy of US-DART. Each vegetation plot corresponded to a growth stage of a homogeneous maize canopy (Duthoit et al., 2008).

(2) **Urban 3D mock-up data**.

This study selected two European cities: Basel, Switzerland, and Brussels, Belgium. Basel (47°33′17″, 7°35′26″), situated in the Rhine Valley, is surrounded by hills and mountains (Feigenwinter et al., 2017; Feigenwinter et al., 2018). A pre-existing 3D mock-up of Basel (Landier, 2018; Landier et al., 2016), including trees and the Digital Elevation Model (DEM), was utilized (Fig. 3). Four scene elements were distinguished: ground, water, building, and vegetation. Trees were simulated using regular octahedrons for trunks and ellipsoidal crowns, uniformly filled with small triangles to represent a turbid medium. However, due to administrative boundaries, the city mock-up and trees were absent in the southwestern part of the study area.

Brussels, located in the north-central part of Belgium (50°51′0″ N, 4°21′0″ E), is approximately 180 km from the country's southern tip. The city, situated on a plateau intersected by various rivers and valleys, has central boulevards elevated 15 m above sea level (https://www.brussels.com/v/geography/). A 3D mock-up of the city was downloaded from Urbis (https://datastore.brussels/web/urbis-download), converted from Sketch-Up format to OBJ format, and consolidated into a single OBJ file. The conversion process encountered difficulties in preserving the classification of the meshes, which were overcome by inferring classes from color information (Fig. 4). The 3D mock-up exhibited certain flaws, such as misclassified roofs and gaps in the DEM, which were rectified using Python scripts and Blender software. The geometric database underwent conversion from the national georeference system, EPSG:31370 (Belgian Lambert 72), to the international system, EPSG:32361 World Geodetic System (WGS) 84 / Universal Transverse Mercator (UTM) zone 31 N. The final OBJ, spanning an area of approximately 16.5 × 16.5 km and comprising over 100 million triangles, was supplemented with tree geometry data furnished by Vrije Universiteit Brussel. This data, generated using an artificial intelligence-based model, provided the position, height, and diameter of the crown and trunk for each of the more than 1.4 million trees. However, due to administrative issues, tree data for the three small rectangle areas (around 4.366°E, 50.887°N; 4.324°E, 50.865°N; and 4.458°E, 50.848°N) were absent (Fig. 4).

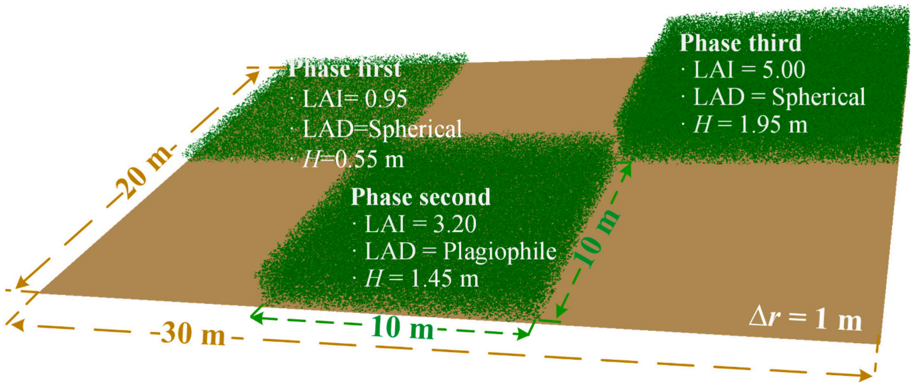


Fig. 2. The vegetation scene: three homogeneous vegetation plots and three bare ground plots. The vegetation parameters, such as leaf area index (LAI), leaf angle distribution (LAD), and canopy height (H), correspond to three development periods of a maize canopy. The LAI is defined per vegetation plot and not for the whole scene. Δr: spatial resolution of the simulated image.

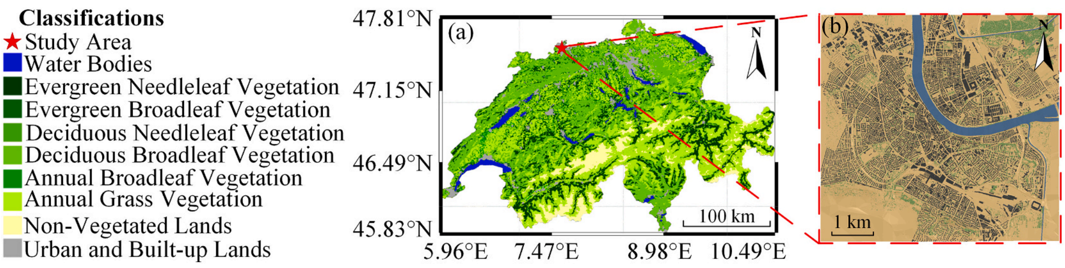


Fig. 3. Study area of Basel. (a) Moderate Resolution Imaging Spectroradiometer (MODIS) land cover type yearly map of Switzerland in 2021. (b) 3D mock-up of Basel. The zone in the southwest appears to be bare ground for administrative reasons.

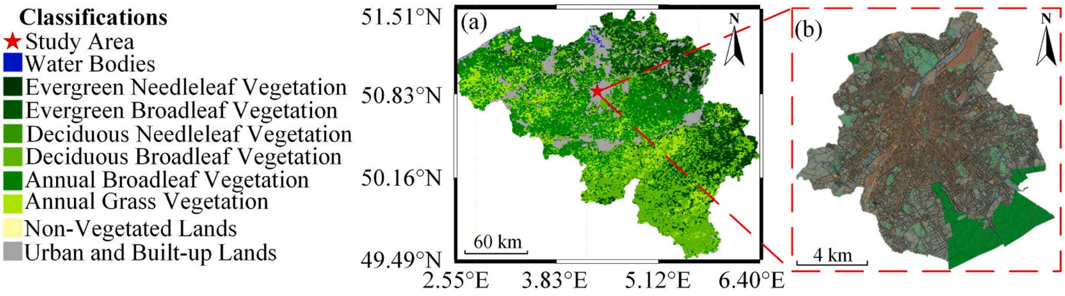


Fig. 4. Study area of Brussels. (a) Moderate Resolution Imaging Spectroradiometer (MODIS) land cover type yearly map of Belgium in 2021. (b) 3D mock-up of Brussels.

2.3.2. *Optical properties of elements*

The OP library, sourced from the DART database, was employed to generate simulated pseudo satellite optical images (Fig. 5). The elements were classified into two categories: transparent (vegetation) and opaque (ground, buildings, and water). Albedo was denoted as OP for transparent elements, while it represented reflectance for opaque elements. Water was treated as an opaque element for simplification. In the simulation of our pseudo satellite images, the OPs for each element were maintained as spatially invariant.

The ratio of reflectance to transmittance in vegetation was subject to considerable variation, influenced by factors such as time, geographical location, and wavelength. Accurately determining this ratio across extensive areas for all vegetation types presented a significant challenge. Therefore, for the cases in Basel, we assumed that reflectance was equal to transmittance and was wavelength-independent. For the cases in Brussels, we relied on empirical values, which were the ratio of the spectrum of deciduous trees from the DART optical database.

2.3.3. *Satellite data and preprocessing*

Orthorectified surface reflectance images from Sentinel-2 L2A and PlanetScope were utilized in this study. The Sentinel-2 images were processed using US-DART, and the resulting element OP maps were validated against PlanetScope images. Sentinel-2 is a high-resolution, multispectral imaging mission with a wide swath, capturing the visible and NIR bands at a fine spatial resolution of 10 m, the red edge and shortwave infrared (SWIR) bands at a moderate resolution of 20 m, and the atmospheric bands at a coarser spatial resolution of 60 m. The B10 (60 m) band was excluded from use as it is primarily used for atmospheric correction and was not included in the L2A collections. These distinct bands were employed to evaluate the sensitivity of US-DART to spatial resolution. PlanetScope, provided by Planet company (https://www.planet.com/), is an imager with eight bands and a 3 m spatial resolution frame, featuring a split-frame visible and NIR filter. Planet offers high-frequency, global coverage, and efficient delivery through an open, accessible platform.

For validation purposes, we selected a pair of images from Sentinel-2 and PlanetScope, taken within a short time frame for both Basel and Brussels, to ensure minimal changes in the OP of land cover over a brief period. Our selection prioritized PlanetScope images with eight bands, followed by those with the shortest interval between Sentinel-2 and PlanetScope captures. Table 1 presents detailed information on the Sentinel-2 and PlanetScope images used for Basel, while Table 2 offers similar details for the images used for Brussels. Table 3 provides a comparison of the spectral and spatial characteristics of the bands between Sentinel-2 and PlanetScope images.

We used image hash similarity and absolute mean difference (AMD) to check the similarities and differences between corresponding bands of Sentinel-2 and PlanetScope. The AMD calculated the absolute difference between corresponding pixel values in two images, then computed the mean of these absolute differences, providing a numerical value representing the average difference between the pixels of the two images. On the other hand, image hashing generated a compact representation (hash) of an image based on its visual features. These hashes were designed to be similar for visually similar images, focusing on capturing the perceptual content of an image rather than exact pixel values. This study used five types of hashes: average, perceptive, difference, Haar wavelet, and Daubechies wavelet. While AMD did not consider the perceptual similarity of images and focused solely on the difference in the overall tone of the image, hashing captured the visual similarity between images, paying more attention to the details and the grey scale difference between pixels (Mahmoud et al., 2007).

Fig. 6 and Fig. 7 show the hash similarity, AMD values, and pixel reflectance distributions for the shared bands in Basel Sentinel-2 and PlanetScope images. The red edge 4 band was of particular interest. It exhibited the highest AMD value (0.0134), indicating a significant overall tonal difference, yet it also had the highest sum of similarity (223.00), suggesting minor differences in detail and grey scale. The median reflectance (0.1784 for Sentinel-2 and 0.1903 for PlanetScope) of the red edge 4 band surpassed that of the other six bands in both satellites (as shown in Fig. 7), leading to the largest AMD. The spectral bandwidth of PlanetScope (0.040 μm) was twice that of Sentinel-2 (0.020 μm), which resulted in lower overall reflectance (median value difference was 0.0119) for the red edge 4 band of Sentinel-2 compared to PlanetScope. However, the other bands' reflectance distribution ranges were similar (median value difference less than 0.003). The hash difference remained minimal despite these differences, resulting in a high similarity.

Fig. 8 illustrates the per-pixel residual error between the Basel Sentinel-2 and PlanetScope images. Given the differing spatial resolutions of the two satellite images (PlanetScope's resolution is 3 m, while Sentinel-2 offers resolutions of 10, 20, and 60 m), we adjusted the Sentinel-2 image to match the column and row count of the PlanetScope image by duplicating pixels. Consistent with the findings in Fig. 6 and Fig. 7, the magnitude of the residual errors were high (mean: 0.032 and median: −0.013) in the red edge 4 band due to its high reflectances than the other bands. Additionally, both images exhibited a significant specular reflection effect (the absolute minimum and maximum values were larger than 1 while reflectance values typically range from 0 to 1).

Fig. 9 and Fig. 10 illustrate the similarities in hashes and AMD values, as well as the pixel reflectance distributions in the shared bands of Brussels Sentinel-2 and PlanetScope images. The Sentinel-2 image exhibited significant over-atmospheric correction, evidenced by the highest AMD value occurring in the aerosol band (0.0173), followed by the blue band (0.0169), indicating a substantial overall tonal difference in these bands. This pattern differed from that observed in Basel (Fig. 6). Despite this, the sum of the similarities in the aerosol (94) and blue (80) bands was not low (Fig. 9), suggesting only minor differences in detail and grey scale. Moreover, the reflectances (median: 0.0258) in the aerosol band in PlanetScope were nearly double those (median: 0.0078) in Sentinel-2 (Fig. 10). Specifically, the pixel reflectances in the Sentinel-2 aerosol band were primarily within the range of 0.00 to 0.02, while in the PlanetScope image, they were mainly within the range of 0.02 to 0.04. For the remaining bands, excluding the aerosol and blue bands, the range of pixel reflectance distributions in the Sentinel-2 and PlanetScope images was nearly identical (difference no larger than 0.0052).

Fig. 11 presents the residual error between the Brussels Sentinel-2 and PlanetScope images. Given the differing spatial resolutions of the two satellite images (PlanetScope had a resolution of 3 m, while Sentinel-2 had resolutions of 10, 20, and 60 m), we adjusted the Sentinel-2 image to match the column and row count of the PlanetScope image by duplicating pixels. As observed in Fig. 9 and Fig. 10, the magnitude of the residual errors were high in the aerosol (mean: 0.020; median: −0.016) and blue (mean: 0.021; median: −0.025) bands compared with other bands, despite the relatively low reflectances in these two bands (median reflectance of S2 equaling 0.0078, of PlanetScope equaling 0.0258 in aerosols band, median reflectance of S2 equaling 0.0169, of PlanetScope equaling 0.0377 in blue band). Additionally, both images exhibited a noticeable specular reflection effect (the absolute minimum and maximum values were larger than 1 while reflectance values typically range from 0 to 1).

The preprocessing phase of the satellite imagery encompassed two critical steps: reprojection and co-registration. To preserve the integrity of the spectral information and minimize distortions induced by resampling, we opted to project the urban 3D mock-up to align with the coordinate systems of the satellite images. Specifically, we utilized the WGS 1984 UTM Zone 32 N for Basel and WGS 1984 UTM Zone 31 N for Brussels. This approach ensured spatial consistency between the 3D model and the remotely sensed data, facilitating accurate analysis and interpretation.

The GeFolki co-registration software (Brigot et al., 2016; Plyer et al., 2015) (https://github.com/aplyer/gefolki), developed by the French Aerospace Lab (ONERA) and achieving a theoretical geometric accuracy of 0.1 pixels, was employed for co-registration between Sentinel-2 and DART images. Any unavoidable inaccuracies in the co-registration of the satellite and DART images could potentially degrade the accuracy of US-DART. However, this degradation could be mitigated if the co-registered images were resampled to a spatial resolution coarser than that of the satellite image. US-DART was capable of performing this spatial resolution degradation automatically using the user's expected value.

3. Application and accuracy assessment of US-DART

Section 3.1 presented an application on vegetation, and made a comparison with DART calibration in terms of inversion accuracy and

Table 1
Parameters of Sentinel-2 and PlanetScope imagery for Basel. Atmosphere data are from the National Centers for Environmental Prediction Reanalysis (NCEP_RE). Atmosphere mode is selected based on the water vapor amount (Cooley et al., 2002). The aerosol mode is selected based on the study target.

Parameters	Values	
Satellite	Sentinel-2	PlanetScope
ID	20211001T102739_20211001T102820_T32TLT	20211001_093505_92_220b_3B_AnalyticMS_8b
Date	2021-10-01 10:27:39 (UTC)	2021-10-01 09:35:05 (UTC)
Mean sun angle (°)	Zenith angle: 51.6, Azimuth angle: 166.1	Zenith angle: 55.5, Azimuth angle: 148.4
Atmosphere model	Subarctic summer (total column water vapor: 14.5 kg/m^2, air temperature: 281.2 K)	
Aerosol model	Urban	

Table 2
Parameters of Sentinel-2 and PlanetScope imagery for Brussels. Atmosphere data are from the National Centers for Environmental Prediction Reanalysis (NCEP_RE). Atmosphere mode is selected based on the water vapor amount (Cooley et al., 2002). The aerosol mode is selected based on the study target.

Parameters	Values	
Satellite	Sentinel-2	PlanetScope
ID	20210224T105031_20210224T105442_T31UES	20210224_100339_42_2276_3B_AnalyticMS_8b 20210224_100341_88_2276_3B_AnalyticMS_8b
Date	2021.02.24 10:50:31 (UTC)	2021-02-24 10:03:39 (UTC) 2021-02-24 10:03:41 (UTC)
Mean sun angle (°)	Zenith angle: 61.7, Azimuth angle: 162.8	Zenith angle: 64.8, Azimuth angle: 149.2
Atmosphere model	Subarctic summer (total column water vapor: 16.7 kg/m^2, air temperature: 284.3 K)	
Aerosol model	Urban	

Table 3
Spectral and spatial comparisons of bands between Sentinel-2 and PlanetScope images.

No.	Des.	Satellites					
		Sentinel-2			PlanetScope		
		λ (μm)	$\Delta\lambda$ (μm)	Δr (m)	λ (μm)	$\Delta\lambda$ (μm)	Δr (m)
B1	Aerosols	0.443	0.020	60	0.442	0.021	3
B2	Blue	0.490	0.065	10	0.490	0.050	3
B3	Green	0.560	0.035	10	0.565	0.036	3
B4	Red	0.665	0.030	10	0.665	0.030	3
B5	RE 1	0.705	0.015	20	0.705	0.016	3
B6	RE 2	0.740	0.015	20	–	–	–
B7	RE 3	0.783	0.020	20	–	–	–
B8	NIR	0.842	0.115	10	–	–	–
B8A	RE 4	0.865	0.020	20	0.865	0.040	3
B9	WV	0.945	0.020	60	–	–	–
B11	SWIR 1	1.610	0.090	20	–	–	–
B12	SWIR 2	2.190	0.180	20	–	–	–

Data are from the Sentinel-2 website (https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric) and PlanetScope product specifications (Marta, 2020). λ: central wavelength, $\Delta\lambda$: spectral bandwidth, Δr: spatial resolution, NIR: near-infrared, SWIR: shortwave infrared, –: no PlanetScope band corresponding to Sentinel-2. RE: Red edge. WV: Water vapor.

efficiency. Section 3.2 presented two applications on urban, involving two European cities, namely Basel in Switzerland and Brussels in Belgium.

3.1. Applications and accuracy assessment of vegetation scene

Section 3.1.1 presented a visual accuracy assessment using a pseudo-satellite vegetation image simulated by DART, accompanied by the display of OP maps of elements and reflectance images. Section 3.1.2 provided a quantitative accuracy assessment of element OP maps and reflectance images. During these assessments, the OP values used to simulate the pseudo-satellite images were considered true values, while the retrieved OP maps from the pseudo-satellite images were treated as estimated values.

3.1.1. Visual accuracy assessment on vegetation scene

We utilized true RGB color composites (Fig. 12) for a visual comparison between the pseudo-satellite image (Fig. 12. a) and DART reflectance (Fig. 12. b) image simulated using the retrieved OP. For instance, the three bare ground plots displayed consistent spatial color, while the three vegetation plots highlighted distinct colors due to their unique leaf area index (LAI) and leaf angle distribution (LAD) in each image. Except that, the pseudo-satellite and DART reflectance images bore a strong resemblance with minimal differences.

Fig. 13 presents the RGB true color composition of the retrieved OP

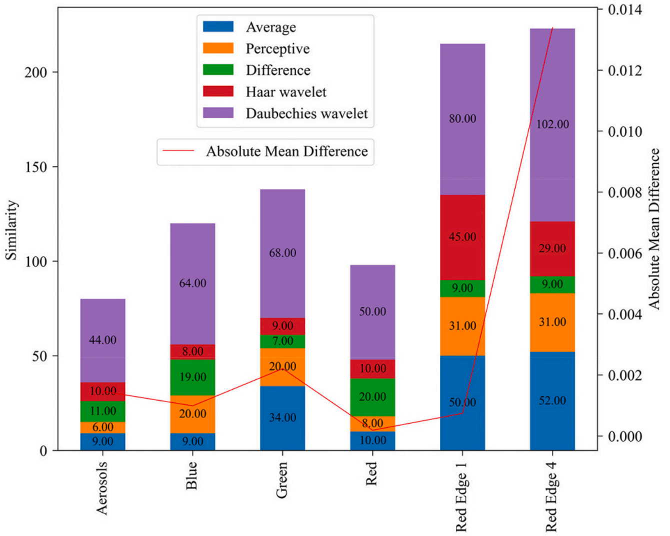


Fig. 6. The hashes similarity (left y-axis) and absolute mean difference (right y-axis) between Basel Sentinel-2 and PlanetScope images in the shared bands.

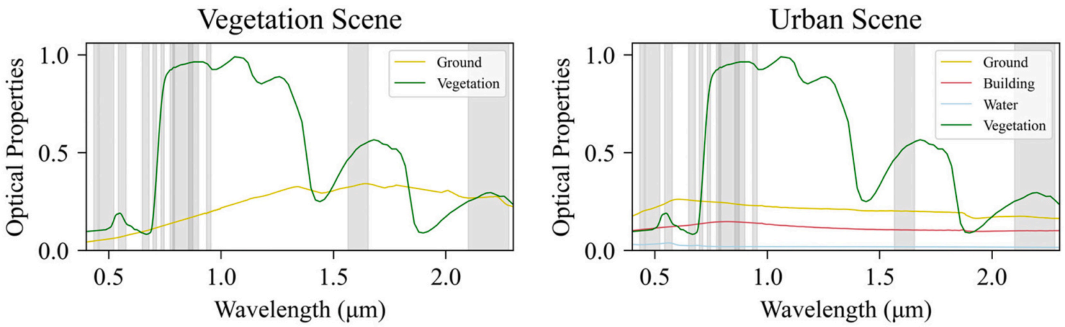


Fig. 5. The optical properties (OP) used for the pseudo satellite image simulation in the vegetation and urban scene. The band configuration is set the same as Sentinel-2. The shade areas indicate the bandwidths of Sentinel-2.

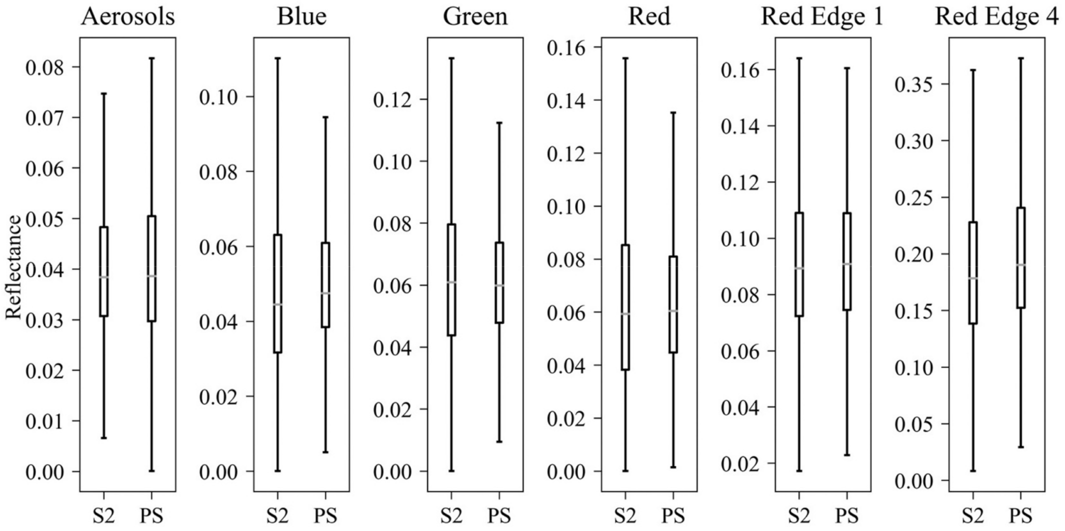


Fig. 7. The reflectance distribution of the shared bands of Basel Sentinel-2 (S2) and PlanetScope (PS) images. Due to strong specular reflection, the image has some outliers, which are not plotted for clarity.

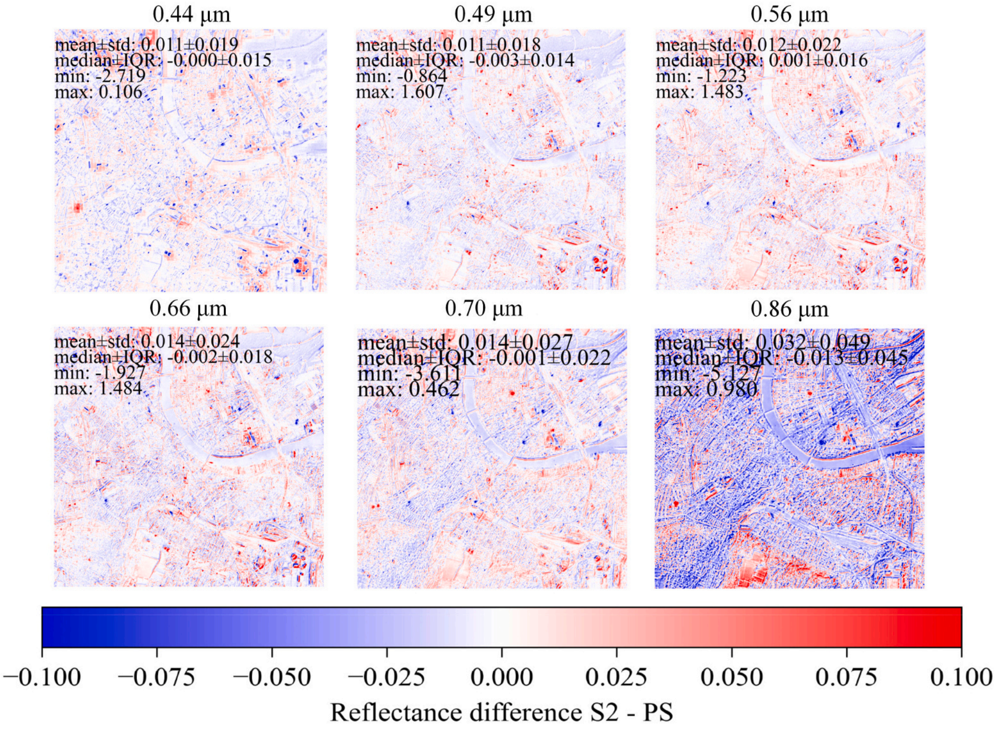


Fig. 8. Residual error of the shared bands between Sentinel-2 (S2) and PlanetScope (PS) of Basel images.

map. The acquired imagery exhibited notable spatial discontinuities. Specifically, the ground pixels (Fig. 13. a) underlying the vegetation canopy displayed significant noise artifacts. Similarly, vegetation pixels (Fig. 13. b) in plots with low LAI demonstrated pronounced noise, particularly during the early stages of plant development. These observations suggested potential limitations in accurately retrieving OP from pixels with low fractional vegetation cover, presumably due to the weak signal in these pixels. Nevertheless, the simulation of pixel reflectance appeared largely unaffected by these high-error OPs, presumably due to the minimal ground fraction within the pixels.

3.1.2. Quantitative accuracy assessment on vegetation scene

To quantitatively evaluate the accuracy of US-DART on vegetation, we utilized the absolute values of the relative error of OP maps and reflectance images. The error variation with iterations was depicted using two statistical parameters: mean ± standard deviation and median ± interquartile range (Fig. 14).

Although the reflectance image error consistently decreased with iterations, the OP map error tended to fluctuate slightly. This fluctuation was evident in the standard deviation and interquartile range, which indicated that the OP map error of some ground pixels experienced several divergences after the PPS module. This divergence occurred because the mixed pixels directly obtained the ground OPs from their nearby flat pure ground pixels in the PPS module. The accuracy of the ground OP from the pure pixels was high due to the predominance of single scattering. However, during the iteration in the GI module, the

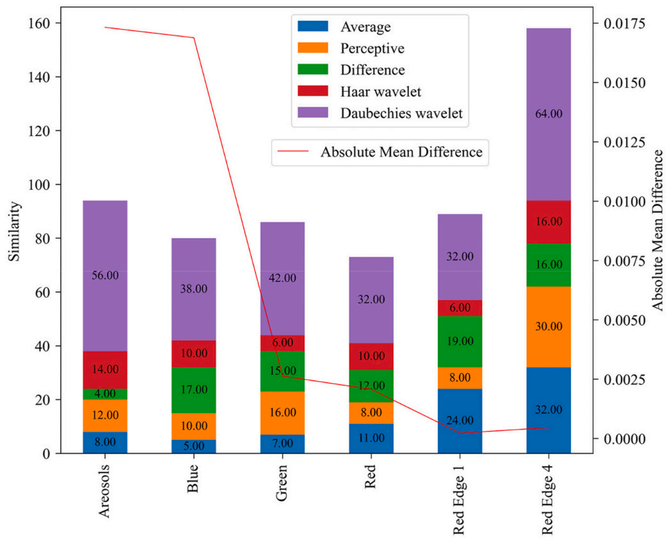


Fig. 9. The hashes similarity (left y-axis) and absolute mean difference (right y-axis) between Brussels Sentinel-2 and PlanetScope images in the shared bands.

ground OP of the very low fraction pixel progressively deteriorated, leading to a slight divergence. In contrast, the vegetation OP error exhibited a clear decreasing trend with iteration in the GI module. This decrease was due to the domain of multiple scatterings in vegetation, resulting in a higher OP error for vegetation from the PPS module than for the ground pixels.

The mean OP error of the ground exhibited minor fluctuations in the bands with longer wavelengths (e.g., longer than the red edge). The reason was that all ground pixels, including those from the low fraction pixels, were taken into account when calculating the mean error. We carefully examined these pixels and found that the majority were ground under high LAI vegetation (e.g., during the third growth period). The median OP error of the ground displayed a decreasing trend with iterations, confirming that only a minimal number of pixels with a low ground percentage demonstrated an increasing error with iteration. In contrast, for the OP error of vegetation, both the mean and median errors consistently decreased with iterations.

At the final iteration, the mean error for reflectance images was approximately 3.47 % for the ground and 5.19 % for vegetation. The median error was around 0.08 % for the ground and 0.20 % for vegetation. For the element OP maps, the mean error was approximately 58.19 % for the ground and 10.95 % for vegetation, while the median error was around 0.60 % for the ground and 3.28 % for vegetation.

3.1.3. Comparison with DART calibration on vegetation scene

Since US-DART is designed to address the challenges of DART calibration caused by mixed pixels, we compare the two methods in terms of accuracy and time complexity. The original DART calibration process exhibits inherent randomness and uncertainty, as it depends on the user's selection of a representative spectral library as the initial iteration values. To ensure consistency and facilitate a more reliable comparison, we integrated the PPS module into DART calibration during the comparison, providing both methods with identical initial values for iteration.

Fig. 15 illustrates the iteration process of DART calibration. The overall accuracy during iteration is significantly lower than that of US-DART. Although certain bands (e.g., RE 4, WV) exhibit a decreasing error trend in pixel reflectance, they show an increasing error trend in OP. At the final iteration, the mean reflectance error was approximately 11.61 % for ground surfaces and 17.61 % for vegetation, while the median error was around 0.37 % and 1.37 %, respectively. For the OP element maps, the mean error reached 126.32 % for ground surfaces and 50.56 % for vegetation, with median errors of 7.40 % and 41.24 %, respectively.

The previous study (Zhen et al., 2021) demonstrates that DART calibration performs optimally in high spatial resolution images (3 m and 4 m) where pure image elements are prevalent. However, in this soil-vegetation mixed scene, we observed suboptimal DART calibration performance. This can be attributed to the Newton and bisection methods implemented in DART calibration, which are designed to account for multiple scattering and generally perform well for pure pixels. However, Eq. (11) in DART calibration (Landier et al., 2018) overlooks the fact that, unlike relative area, spectral contributions vary in conjunction with OP. To address this limitation, we replaced the Newton and bisection methods with the GI module, improving accuracy in scenes dominated by mixed pixels.

We also compared the time complexity of US-DART and DART calibration. Our improvements regarding acceleration compared with the original DART calibration focused on accelerating convergence and optimizing parallel processing. To enhance convergence, we introduced a PPS module that selects initial values from pure pixels rather than relying on priori knowledge (e.g., library values), as in the original DART calibration. This modification reduces the number of iterations

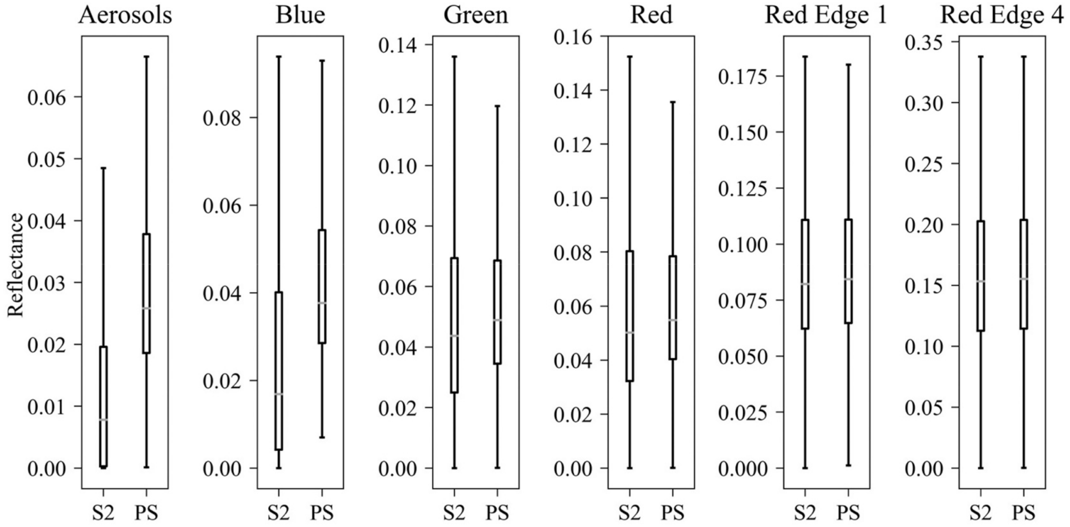


Fig. 10. The reflectance distribution of shared bands of Brussels Sentinel-2 (S2) and PlanetScope (PS) images. Due to strong specular reflection, the image has some outliers, which are not plotted for clarity.

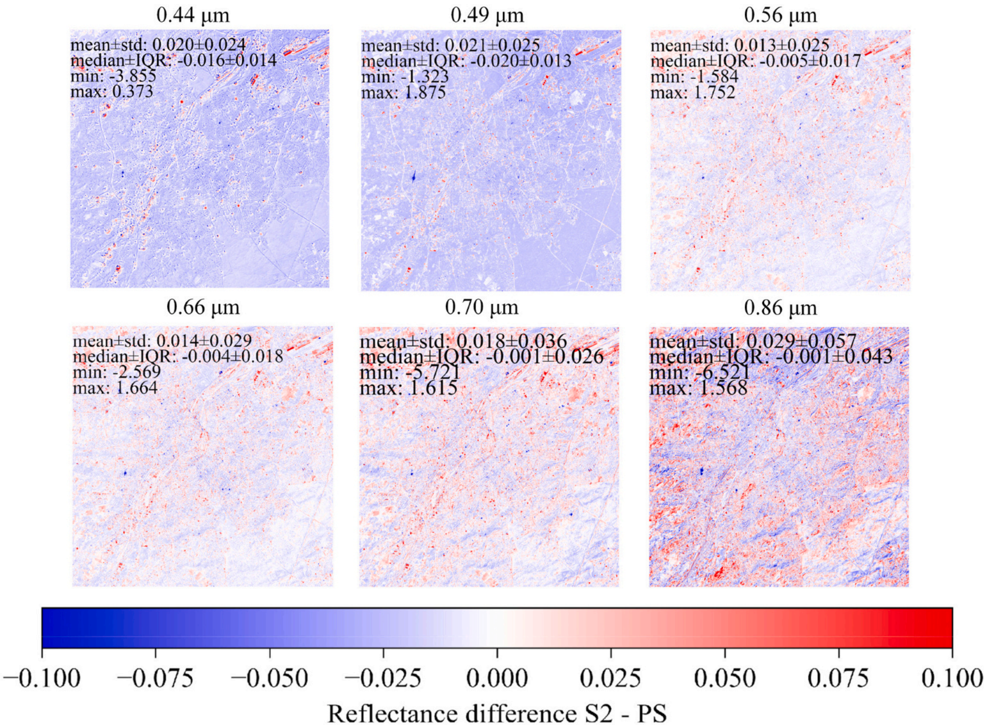
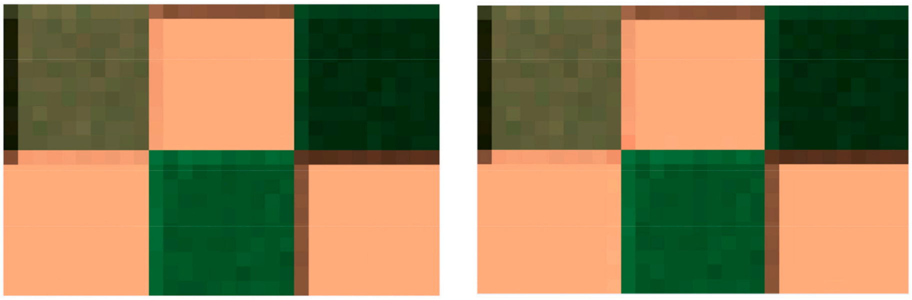
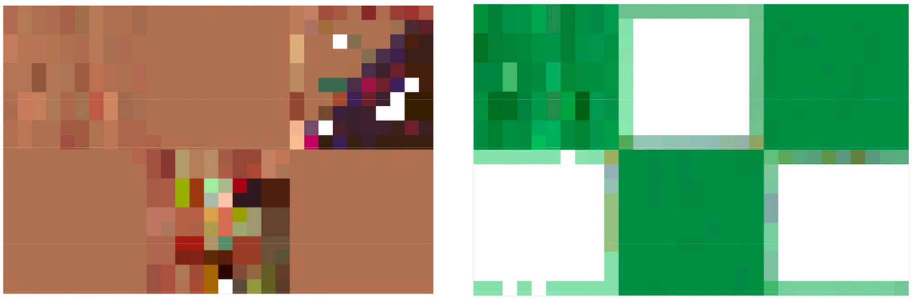


Fig. 11. Residual error of the shared bands between Sentinel-2 (S2) and PlanetScope (PS) of Brussels images.



(a) Pseudo satellite image (b) DART image

Fig. 12. The true RGB composites of the vegetation reflectance image. (a) Pseudo satellite image; (b) DART image simulated using retrieved element OP maps from US-DART.



(a) Ground OP map (b) Vegetation OP map

Fig. 13. The true RGB color composites of the retrieved ground and vegetation optical properties (OP) maps from US-DART. (a) Ground OP map. (b) Vegetation OP map. The white pixels in the image indicate that the fractions of the corresponding elements are too small in these pixels and are ignored.

required. For parallel processing, we implemented multi-threading: for multi-band images, each band is processed independently on a separate thread, while for mono-band images, the image is divided into blocks, each processed on a separate thread. However, to ensure a mathematically rigorous comparison and minimize uncertainty from human intervention, the following DART calibration used in the comparison also incorporated these acceleration techniques.

All experiments were conducted on a workstation equipped with an

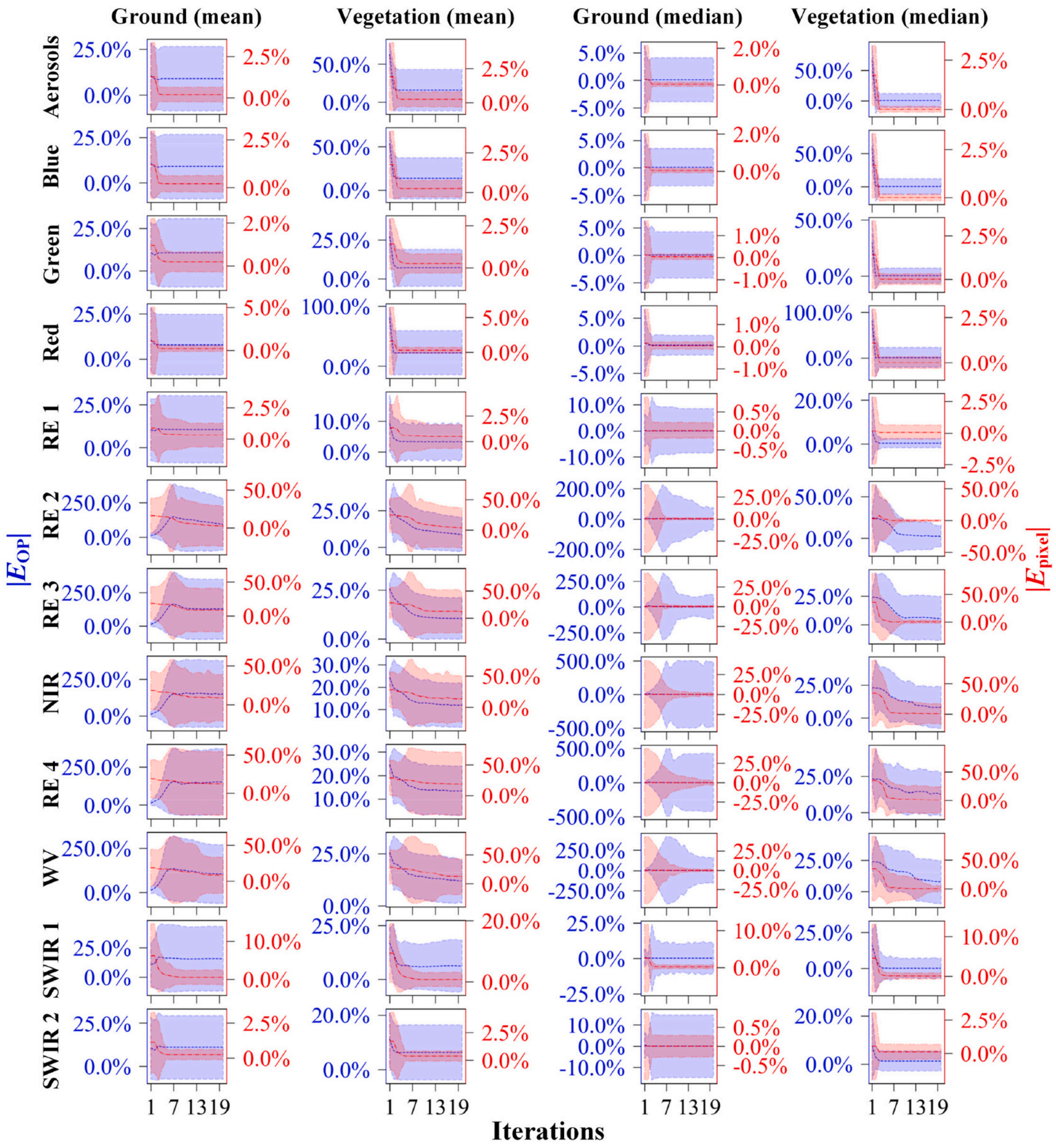


Fig. 14. Quantitative accuracy assessment of US-DART on vegetation scene. The absolute values of relative error of the retrieved ground and vegetation optical properties maps (blue left y-axis) and simulated reflectance images (red right y-axis) using US-DART. RE: Red edge. NIR: Near infrared. WV: Water vapor. SWIR: Shortwave infrared. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Intel Core i7–8700 processor (3.2 GHz, 6 cores, 12 threads) and 62 GB of RAM. The system ran Ubuntu 22.04.2 LTS (Jammy) and utilized a storage configuration comprising a 476.9 GB primary SSD and a 931.5 GB secondary encrypted disk. The implementation was performed in a Linux environment, leveraging multi-threaded processing capabilities. Table 4 presents a temporal comparison between US-DART and DART calibration methods across eight iterations for the vegetation scene, with each iteration comprising two processes: DART simulation and inversion. For DART simulation, both methods demonstrated relatively consistent performance, with execution times ranging between approximately 202–216 s. US-DART's simulation times varied from 204.85 to 215.74 s, while DART calibration maintained slightly more stable timing between 202.59 and 206.88 s.

The most notable distinction between the two methods emerged in their inversion processes. The initial iteration showed comparable inversion times of around 198 s for both methods. However, from iteration 1 onward, the patterns diverged significantly. US-DART exhibited a distinctive spike in iteration 2, with an inversion time of 832.72 s (due to the calculation time consumption of gradient), but subsequently stabilized to approximately 10 s for the remaining iterations. In contrast, DART calibration maintained more consistent inversion times after iteration 0, ranging between 13 and 17 s, though slightly higher than US-DART's stabilized performance. Although US-DART takes more time to calculate gradients, its overall speed is faster once it stabilizes. We are currently working on reducing the time of gradient calculation.

3.2. Applications and accuracy assessment of urban scene

Section 3.2.1 detailed the accuracy assessment conducted using Sentinel-2 images. Section 3.2.2 presented the accuracy assessment performed using pseudo satellite images simulated by DART. During these assessments, the OP inputted to simulate the pseudo satellite image was considered the true value, while the OP retrieved from the pseudo satellite image was regarded as the estimated value.

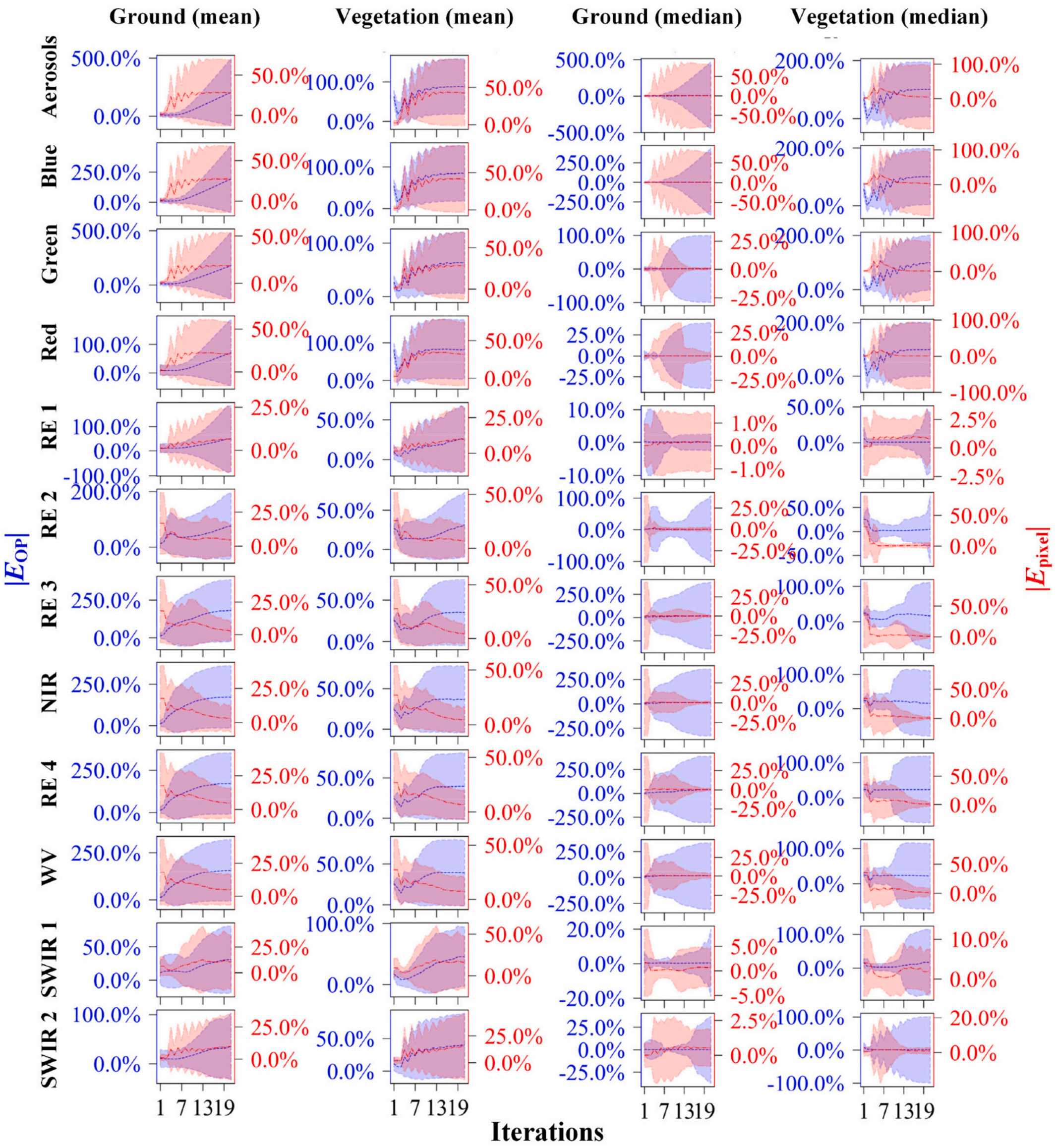


Fig. 15. Quantitative accuracy assessment of DART calibration on vegetation scene. The absolute values of relative error of the retrieved ground and vegetation optical properties maps (blue left y-axis) and simulated reflectance images (red right y-axis) using DART calibration. RE: Red edge. NIR: Near infrared. WV: Water vapor. SWIR: Shortwave infrared. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

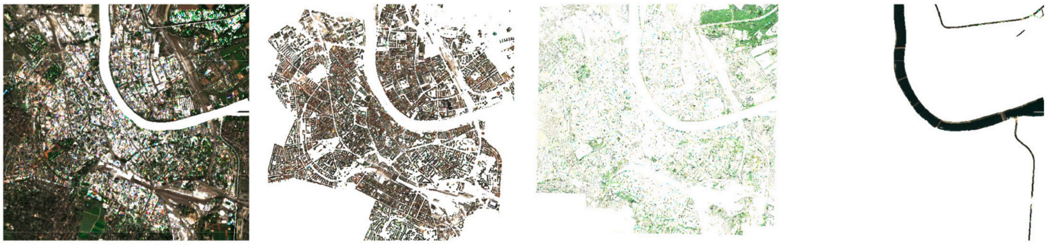

(a) Ground (b) Building (c) Vegetation (d) Water

Fig. 16. The RGB true color composition of the optical properties (OP) maps of the four elements of Basel. (a) Ground. (b) Building. (c) Vegetation. (d) Water. Due to administrative boundaries, the city mock-up and vegetation are missing in the lower left area. The white pixels in the map indicate that the fraction of the corresponding elements is too small in these pixels and is ignored.

3.2.1. Accuracy assessment with Sentinel-2 multispectral instrument (MSI) image

(1) **Visual accuracy assessment on the urban scene.**

Fig. 16 and Fig. 17 display the true RGB color compositions of the OP maps retrieved from Sentinel-2 for Basel and Brussels, respectively. In both study areas, each element's OP exhibited a uniform spatial distribution with minimal noisy pixels, aligning with the geostatistics theory

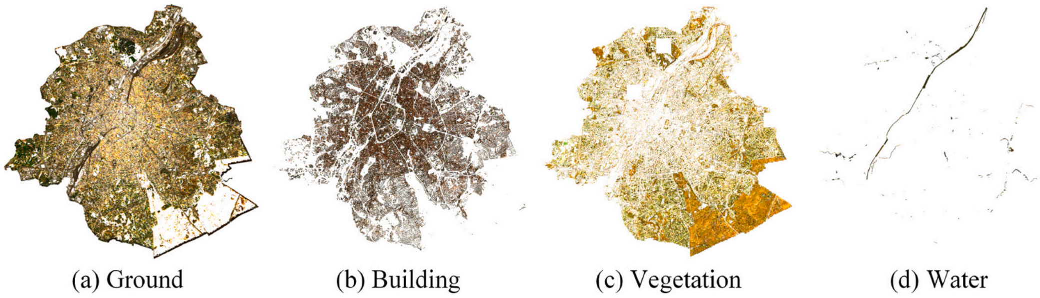


Fig. 17. The RGB true color composition of the optical properties (OP) maps of the four elements of Brussels. (a) Ground. (b) Building. (c) Vegetation. (d) Water. Three rectangular vegetated areas in the map are missing for administrative reasons. The white pixels in the map indicate that the fraction of the corresponding elements is too small in these pixels and is ignored.

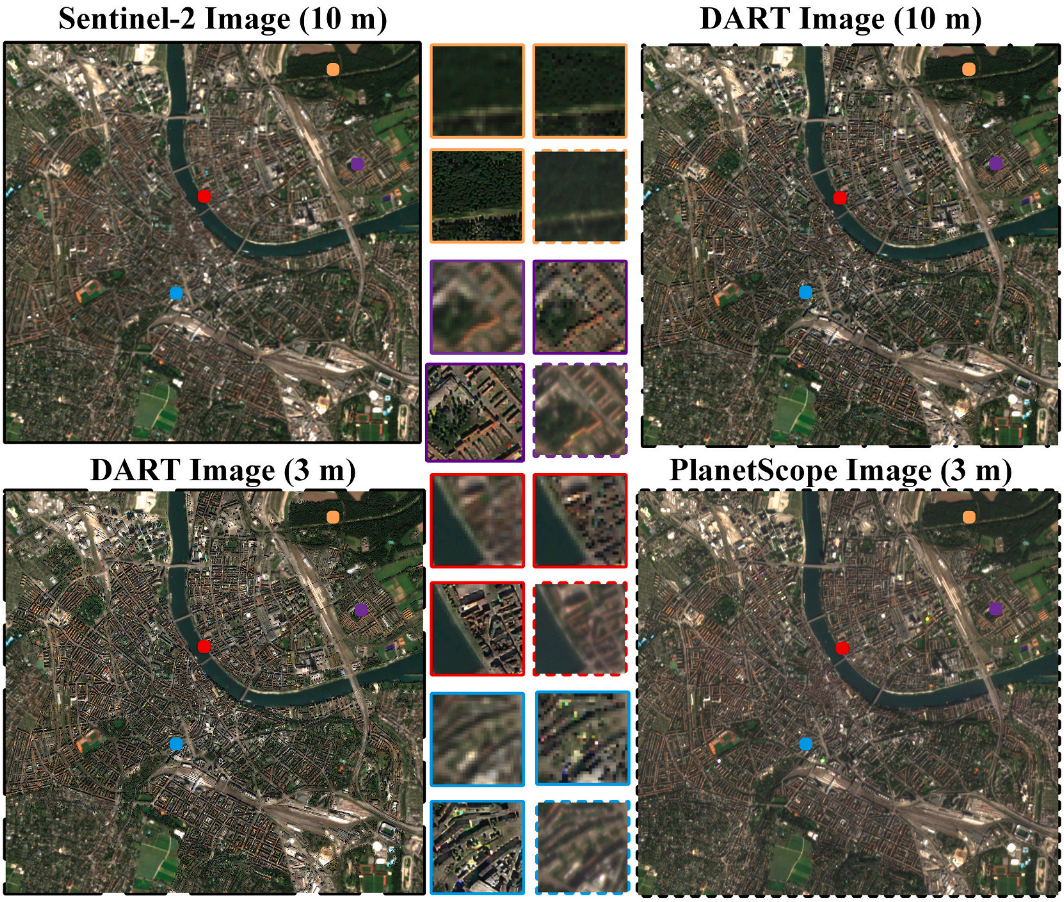


Fig. 18. The RGB true color composition of Basel surface reflectance images. The black frames show the whole study area, and the four colorful frames show the zooms of the subzones. Yellow: vegetation subzone. Purple: building-vegetation mixed subzone. Red: building-water subzone. Blue: built-up subzone. The line styles of the frames distinguish the image types. Solid line: inputted Sentinel-2 images. Dashed and dotted line: DART simulated 10 m resolution images. Solid and dashed line: DART simulated 3 m resolution images. Dashed line: PlanetScope images. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that nearby elements shared similar OPs. Furthermore, each element presented visually distinguishable OP maps, signifying the successful separation of each element's OP from the mixed pixels. A comparison of the OP maps between Basel and Brussels revealed greener vegetation in Basel. This difference was likely attributable to the distinct seasons captured (Basel: Autumn, 1 October; Brussels: Winter, 24 February) and the over-atmospheric correction applied to the Brussels image.

The selection of RGB true-color composite map for our analysis served two key purposes. First, these images aligned with US-DART's novel methodological framework, which inverted the traditional mixed-pixel modeling approach. While conventional models treated endmember spectra as fixed quantities and solve for abundances, US-DART employed 3D radiative transfer to generate spectral gradients that replace abundance values. This transformation enabled endmember spectra to become the primary unknown variables, allowing them to vary on a per-pixel basis within the same component type. Such flexibility was particularly valuable in urban environments, where material properties like rooftop spectra naturally varied due to weathering and material heterogeneity. Second, the RGB true-color composite visualization demonstrated the effectiveness of US-DART's band-by-band spectral retrieval method. Unlike traditional approaches that depended on inter-band correlations, US-DART resolved endmember spectra

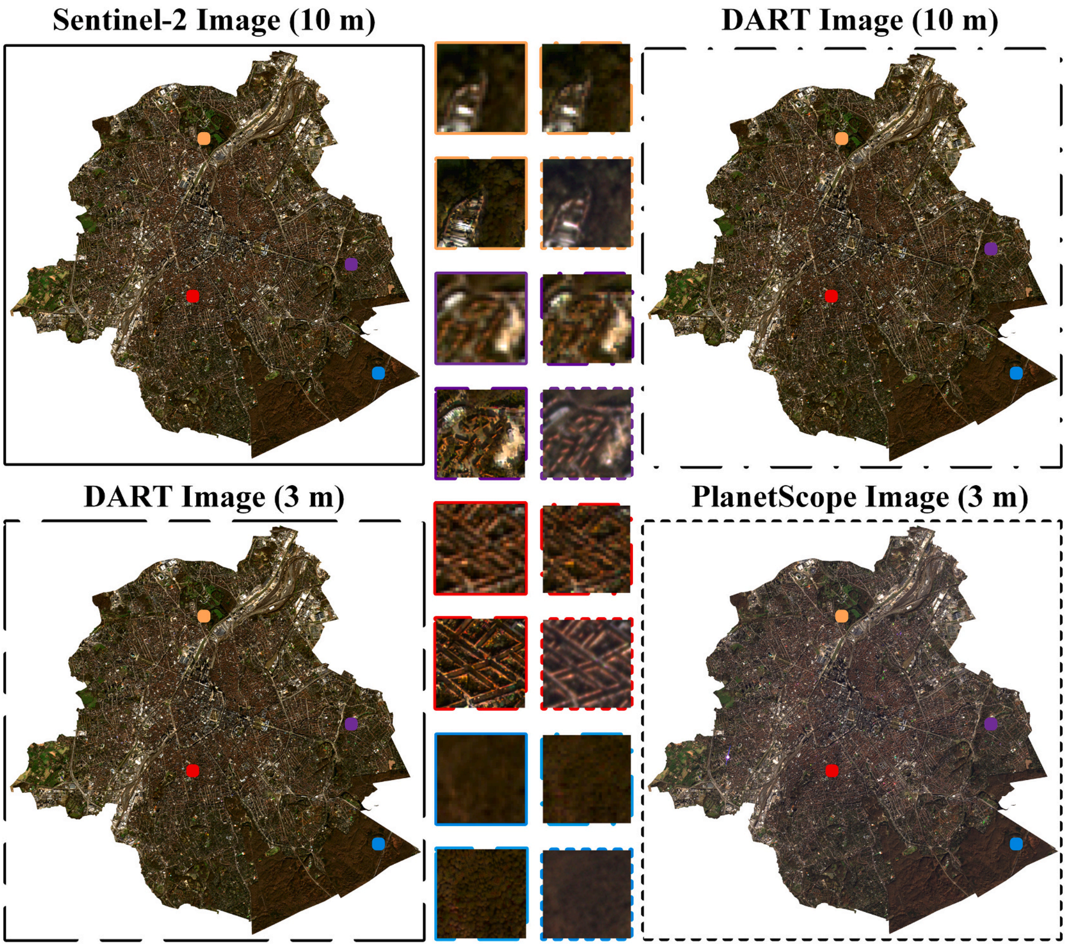


Fig. 19. The RGB true color composition of Brussels surface reflectance images. The black frames show the whole study area, and the four colorful frames show the zooms of the subzones. Yellow: building-vegetation mixed subzone. Purple: built-up sparse subzone. Red: built-up dense subzone. Blue: vegetation subzone. The line styles of the frames distinguish the image types. Solid line: Inputted Sentinel-2 images. Dashed and dotted line: DART simulated 10 m resolution images. Solid and dashed line: DART simulated 3 m resolution images. Dashed line: PlanetScope images. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

independently for each band, making it available for single-band applications. This visualization technique served as a powerful diagnostic tool, as retrieved spectral errors manifested as distinct noise patterns in the composite image.

We simulated a 3 m spatial resolution image using the US-DART retrieved element OP map for Basel and Brussels for further validations. Fig. 18 and Fig. 19 display the RGB true color composition of the inputted Sentinel-2 image, the simulated 10 m resolution image, the simulated 3 m resolution image, and the PlanetScope image for Brussels and Basel, respectively. The visual differences between the two simulated images at 10 m and 3 m resolutions and the Sentinel-2 images were nearly indistinguishable across the study area. The PlanetScope image, however, exhibited a noticeable difference in tone compared to the other three images, likely due to the differences of instruments between PlanetScope and Sentinel-2. For instance, PlanetScope typically exhibits lower radiometric quality than Sentinel-2 and Landsat 8 (Houborg and McCabe, 2018; Myers et al., 2019).

We examined four subzones in Basel (i.e., one vegetation area, one building-vegetation mixed area, one building-water area, and one built-up area) and Brussels (i.e., one building-vegetation mixed area, one built-up sparse area, one built-up dense area, and one vegetation area). The DART simulated images appeared sharper in these subzone images than the satellite images, particularly the DART simulated 10 m resolution images were sharper than the Sentinel-2 images, and the DART simulated 3 m resolution images were sharper than the PlanetScope images. The reason may be that the DART simulated images are free from the sensor noise (e.g., modulation transfer function) and atmospheric scattering.

The DART-simulated images of 3 m and 10 m spatial resolutions exhibited a few noisy pixels, likely attributable to the unavoidable inaccuracies of the urban 3D mock-up and the co-registrations, leading to inaccurate gradient estimation. The noise in the 10-meter resolution image was lighter compared to the 3-meter resolution image because the proportion of components within pixels differed at various resolutions. At low resolution (10 meters), the abundance of OP was low, resulting in poorer retrieval accuracy but less interference with the simulated reflectance image. At high resolution (3 meters), this component was likely to become dominant with increased abundance, leading to noisier pixels when simulating reflectance images. Furthermore, ideal assumptions of the US-DART, such as the same element in the same sampling windows sharing the same OP, were also likely to contribute to the generation of noisy pixels.

(2) **Quantitative accuracy assessment on the urban scene.**

To quantitatively evaluate the accuracy of US-DART, we utilized the relative error for pixel reflectance. Two statistical parameters, mean error ± standard deviation and median error ± interquartile range, were employed to illustrate the error variations with iterations.

Fig. 20 and Fig. 21 show the mean and median error with iterations

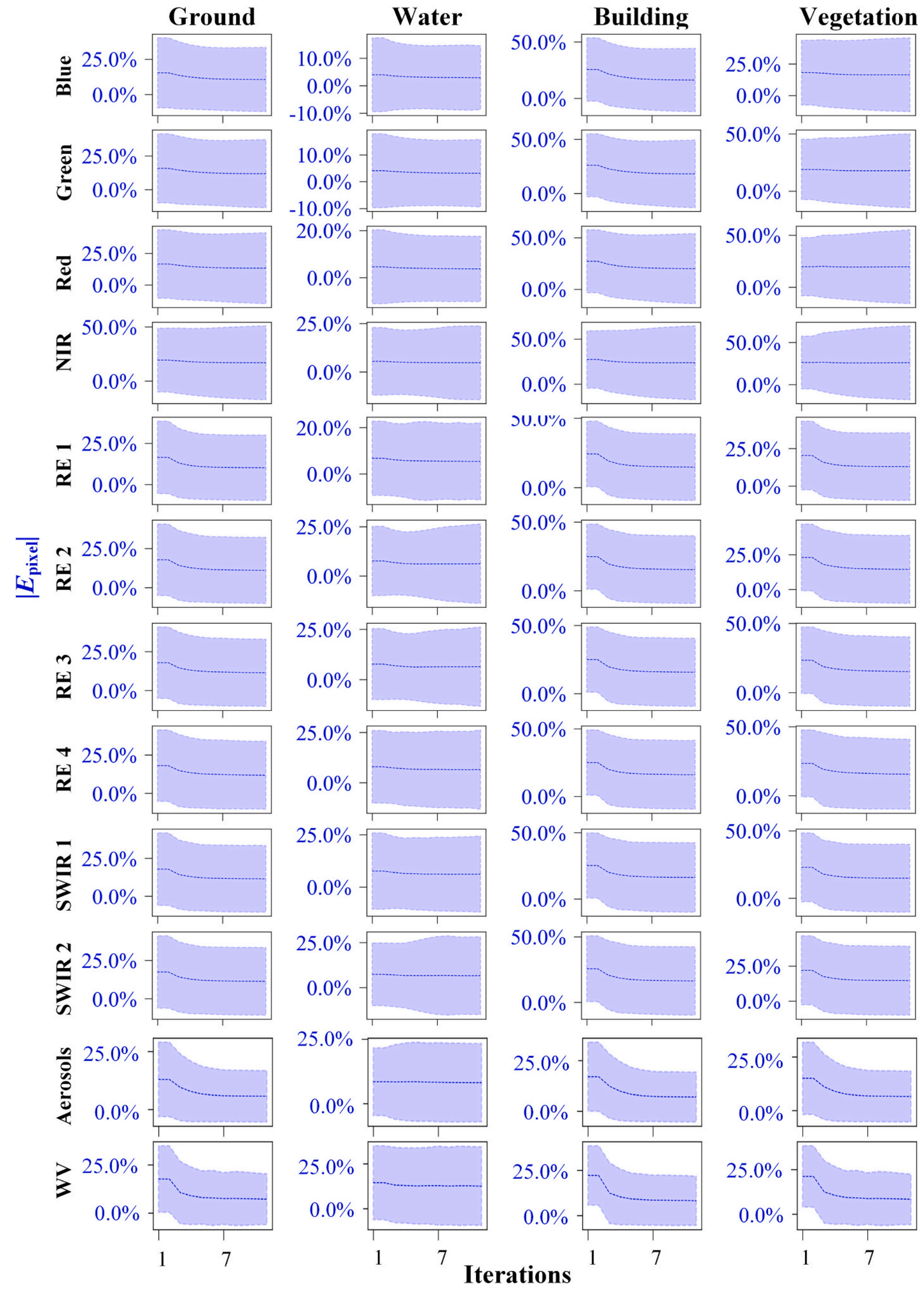


Fig. 20. The variations of the mean error with iterations for the Basel. The absolute values of relative error are calculated with US-DART simulated pixel reflectance as estimated values and Sentinel-2 pixel reflectance as true values. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor.

for the Basel image. All errors decreased with iterations, with the median error showing a more pronounced decreasing trend than the mean error, indicating that most pixels demonstrated promising results with iteration. However, a few pixels exhibited an increase in error, typically due to outdated urban mock-ups or co-registration errors. At the final iteration, the mean errors for ground, water, building, and vegetation were approximately 11.04 %, 6.03 %, 15.60 %, and 15.18 %, respectively. The corresponding median errors were around 0.01 %, 0.04 %, 0.06 %, and 0.04 %.

Fig. 22 and Fig. 23 present the mean and median error with iterations for the Brussels image. Similar to Basel, all errors decreased with iterations, and the median error showed a more noticeable decreasing trend than the mean error, indicating that most pixels demonstrated promising results with iterations. However, a few pixels exhibited an increase in error, typically due to outdated urban mock-ups or co-registration errors. The mean error in the aerosol, blue, and NIR bands was notably high (Fig. 22). Upon examination, we found some pixels with nearly zero reflectance in the forest areas (likely due to atmospheric over-correction), which tended to generate unexpected relative errors. These three bands exhibited normal relative error for median error (Fig. 23). At the final iteration, the mean errors for ground, water, building, and vegetation were approximately 38.07 %, 57.46 %, 38.94 %, and 43.64 %, respectively. The corresponding median errors were around 0.06 %, 0.03 %, 0.09 %, and 0.20%.

3.2.2. *Accuracy assessment with pseudo satellite images simulated by DART*

We utilized DART simulated reflectance images as pseudo satellite images, which provided the advantage of having true element OP values for quantitative assessment of US-DART's accuracy. The assessment

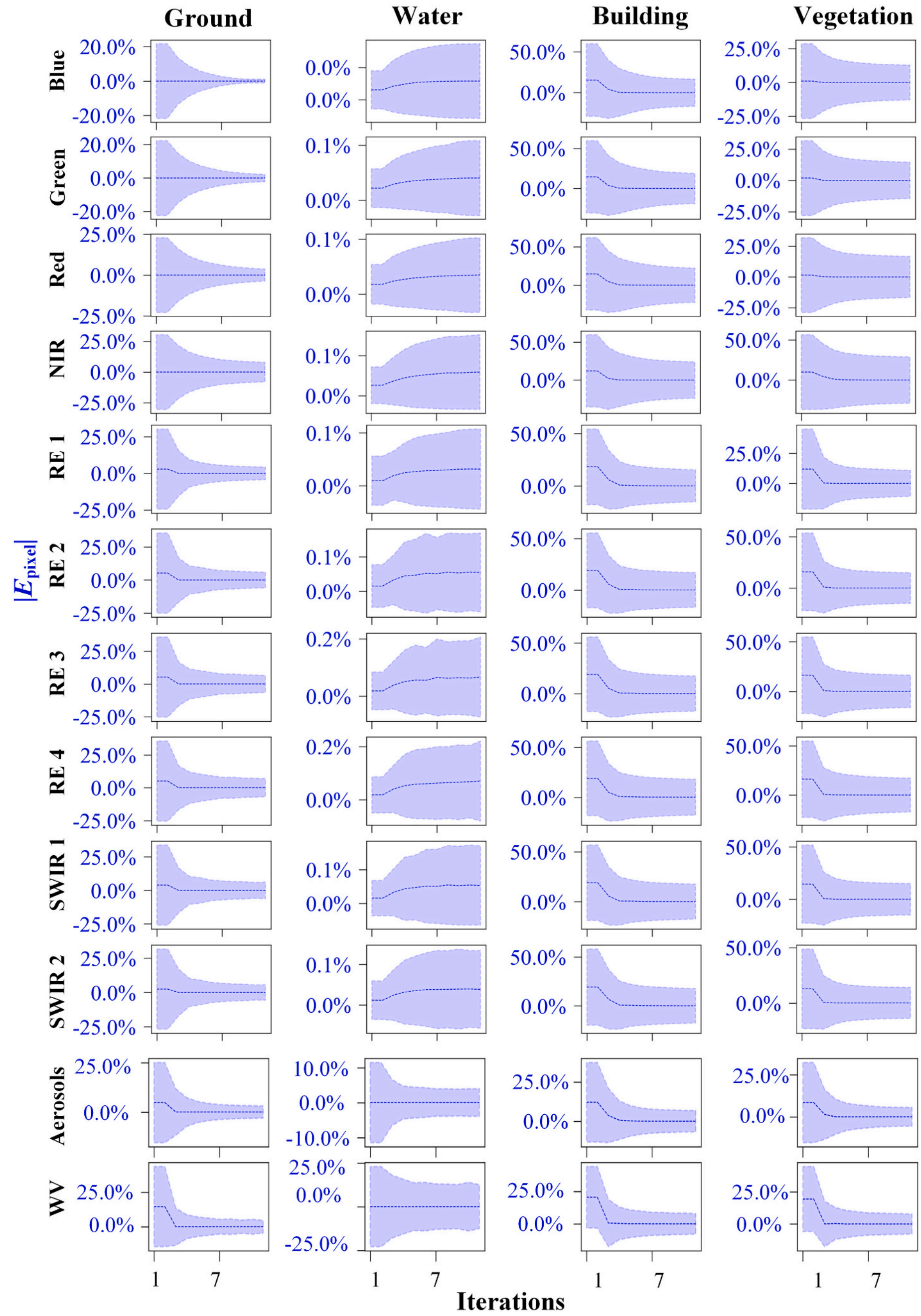


Fig. 21. The variations of the median error with iterations for the Basel. The absolute values of relative error are calculated with US-DART simulated pixel reflectance as estimated values and Sentinel-2 pixel reflectance as true values. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor.

involved both reflectance images and OP maps. To illustrate the variation in relative error with iterations, we employed two statistical parameters: mean error ± standard deviation and median error ± interquartile range.

Fig. 24 and Fig. 25 illustrate the mean and median error trends across iterations for the Basel image. The ground, buildings, and vegetation exhibited a clear decreasing trend in reflectance image and OP map errors with each iteration. Water, however, was an exception, as its error variation trend was unstable and sometimes even showed an increasing trend. This instability was attributed to two main reasons: first, because compared with other components, the surface of rivers is relatively smooth. Therefore, the OP accuracy of water obtained from pure pixels is already very high, making further accuracy improvements challenging; second, due to the low OP of water relative to other elements, the effect of Monte Carlo noise was more pronounced when calculating the gradients of water. Additionally, the median relative error generally showed a much more apparent decreasing trend than the mean relative error, indicating that the error of most pixels decreased with iterations.

At the final iteration, the mean errors for ground, water, building, and vegetation were approximately 8.67 %, 5.06 %, 11.98 %, and 12.62 %, respectively. The corresponding median errors were around 0.01 %, 0.04 %, 0.04 %, and 0.48 %. The pseudo image application demonstrated superior accuracy compared to Sentinel-2 images due to their freedom from co-registration errors, atmospheric interference, and sensor noises. The fluctuation was slightly more pronounced for the OP map error than the reflectance image error, but a clear decreasing trend was evident in almost all lines. At the final iteration, the mean errors for ground, water, building, and vegetation were approximately 16.12 %, 9.99 %, 17.82 %, and 54.40 %, respectively, while the median errors were around 0.80 %, 1.30 %, 5.15 %, and 29.51 %. Multiple

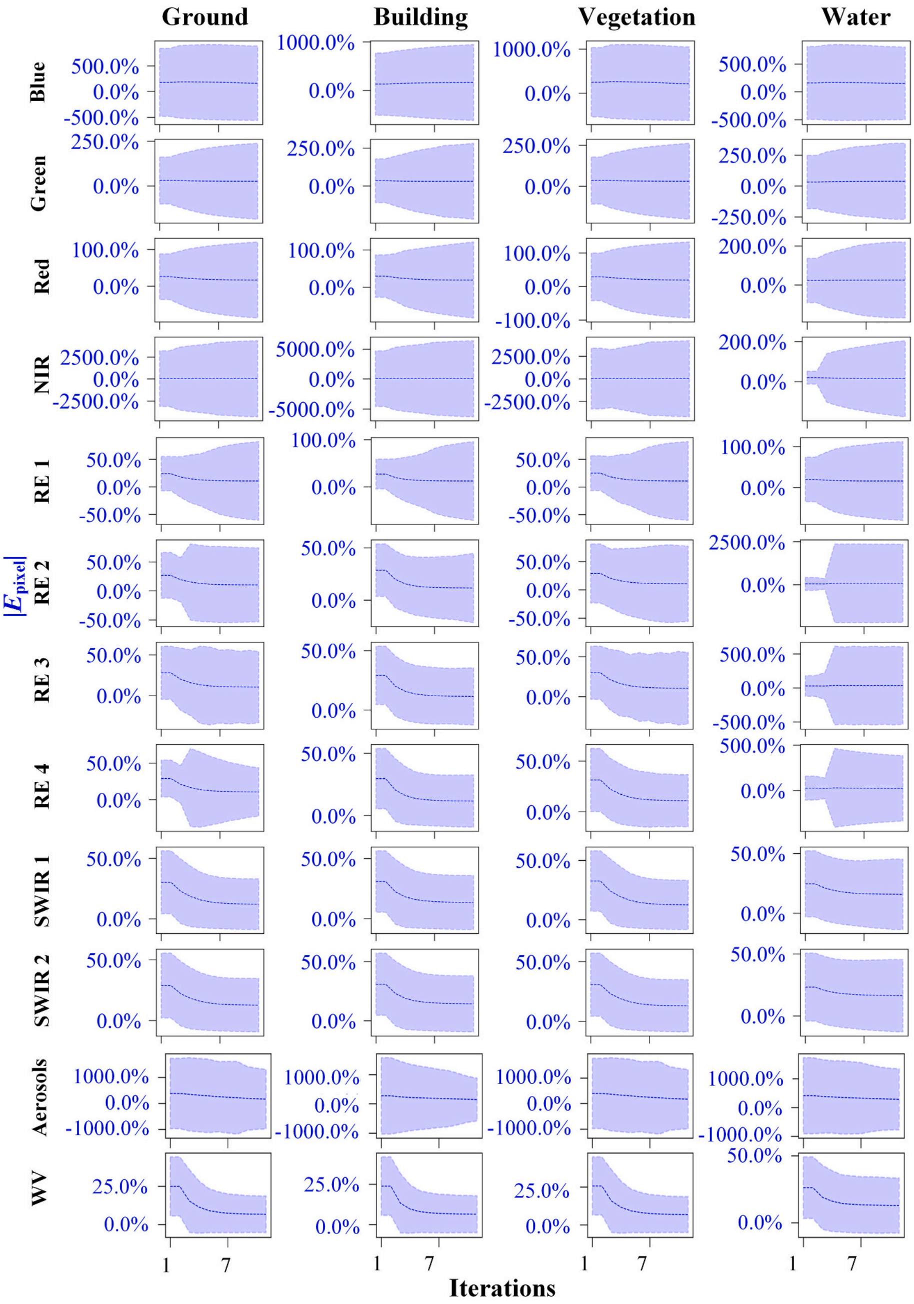


Fig. 22. The variations of the mean error with iterations for the Brussels. The absolute values of relative error are calculated with US-DART simulated pixel reflectance as estimated values and Sentinel-2 pixel reflectance as true values. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor.

combinations of element OPs were able to result in the same pixel reflectance, which typically led to a lower retrieval accuracy for the OP map compared with the reflectance image.

Fig. 26 and Fig. 27 present the mean and median relative error trends across iterations for Brussels, mirroring the decreasing trend observed in the Basel image. At the final iteration, the mean reflectance image errors for ground, water, building, and vegetation were approximately 1.16 %, 1.63 %, 0.75 %, and 1.39 %, respectively, while the median errors were around 0.07 %, 0.05 %, 0.07 %, and 0.06 %. For the OP map error, the mean errors at the final iteration for ground, water, building, and vegetation were approximately 9.23 %, 24.31 %, 3.19 %, and 8.75 %, respectively, with median errors of around 0.28 %, 0.59 %, 0.07 %, and 1.21 %.

The retrieval accuracy of vegetation in Brussels significantly surpassed that in Basel due to the input of the correct leaf reflectance to transmittance ratio. Basel's approach assumed reflectance was equal to transmittance, leading to inaccuracies. This incorrect ratio had a negligible impact on the simulation accuracy of canopy reflectance but significantly affected retrieving leaf albedo. Upon careful examination, we identified the primary errors originating from the aerosol, blue, and red bands. The relatively low vegetation reflectance in these bands resulted in a significant calculated relative error. Furthermore, the disparity between reflectance and transmittance in these bands contributed to additional errors in the assumption that reflectance equaled transmittance.

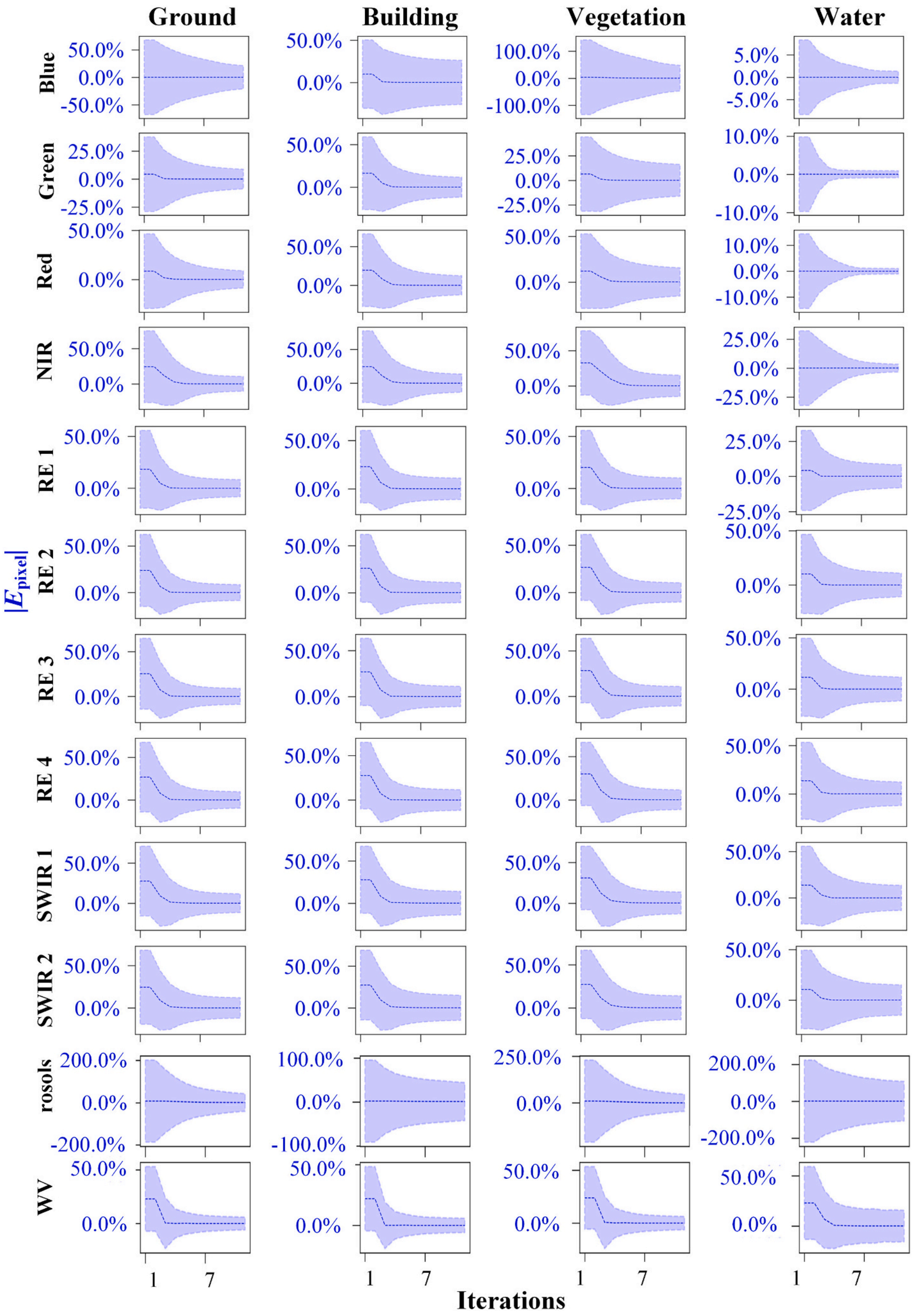


Fig. 23. The variations of the median error with iterations for the Brussels. The absolute values of relative error are calculated with US-DART simulated pixel reflectance as estimated values and Sentinel-2 pixel reflectance as true values. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor.

4. Discussion and conclusions

4.1. Discussion

4.1.1. The advantages of the US-DART for retrieving element optical properties from mixed pixels

US-DART can unmix monospectral images, which capture a wide range of wavelengths in a single band, providing high spatial resolution but limited spectral information (Bay et al., 2010). Additionally, while US-DART requires few input parameters, incorporating rich prior knowledge, such as the OP rules of vegetation, can enhance its accuracy, which is particularly evident in the SC module's utilization of prior knowledge.

US-DART excels in integrating grid points and categories, a focus area of growing research interest. While some studies (Fontanilles and Briottet, 2011) use grid points from DSM for nonlinear spectral unmixing and others (Mitraka et al., 2012) utilize categories from land cover classification maps, the combination of grid points and categories, forming a 3D mock-up is rarely used. US-DART's innovative application of a 3D mock-up for unmixing SW optical images provides nonlinear unmixing conditions and specific categories for each land cover type, thereby enhancing overall effectiveness.

The significant distinction between US-DART and traditional spectral unmixing methods is the output product. Traditional spectral unmixing methods retrieve 2D class fractions from mixed pixels. However, US-DART retrieves pixel-wise endmember spectra from mixed pixels, and the class fractions during unmixing process are replaced by spectral gradients calculated using 3D radiative transfer. The spectral gradients describe the ray intercept and transfer among facets in the 3D mock-up, rather than the relative area of the corresponding endmember

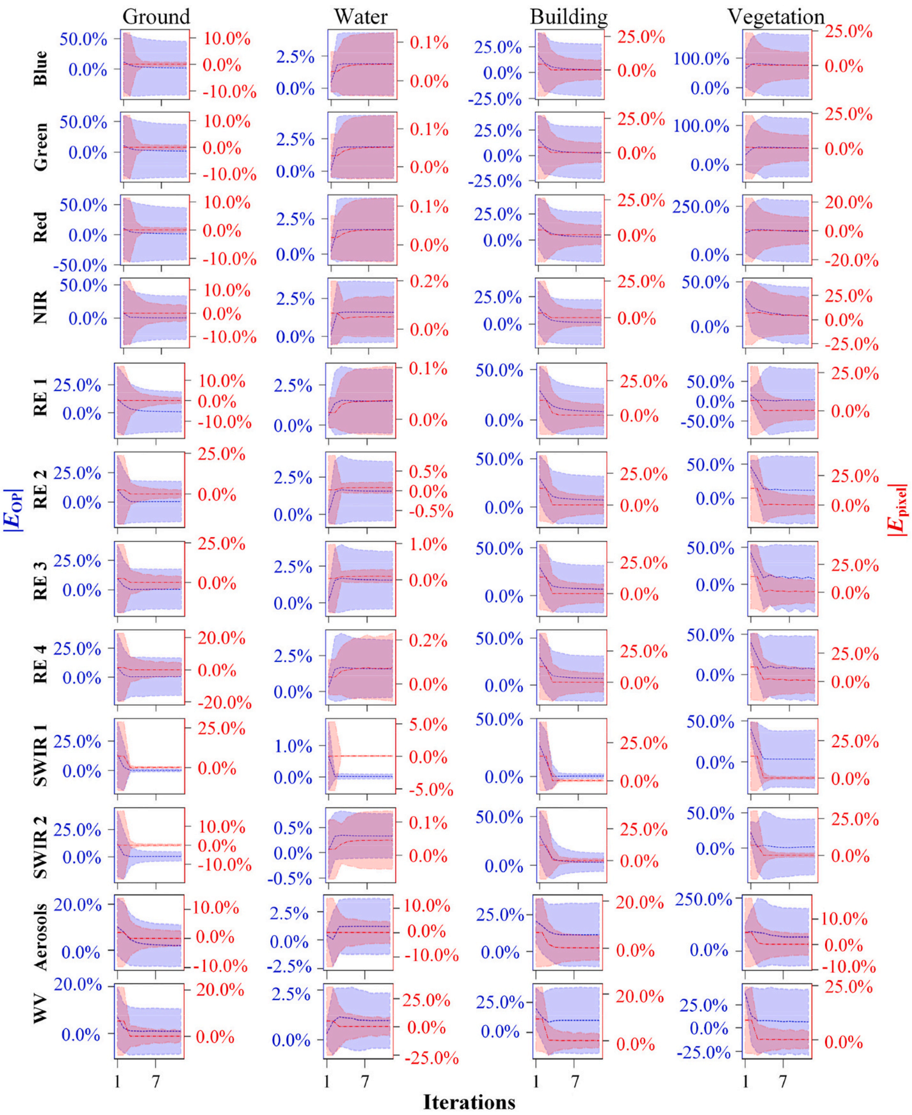


Fig. 24. The mean error variations with iterations for the Basel scene. The left blue y-axis shows the absolute values of relative error of retrieved optical property maps. The right red y-axis shows the absolute values of the relative error of simulated reflectance images. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

used in traditional spectral unmixing methods. Therefore, US-DART allows accounting for factors like vertical structure and material heterogeneity, which traditional 2D spectral unmixing methods like LSMA and Multiple Endmember Spectral Mixture Analysis (MESMA) (Franke et al., 2009) may not fully capture. US-DART holds a distinct advantage over existing methods in addressing multi- and non-endmember problems. US-DART is a "non-endmember" method capable of handling images with a certain level of noise, a task that poses a challenge for most non-endmember methods because these conventional methods typically seek extreme points in the feature space (Drumetz et al., 2019b). US-DART's advantage stems from its use of a 3D mock-up, enabling it to accurately identify endmembers and account for illumination conditions, free from seeking extreme points.

Machine learning advancements in regression-based spectral unmixing have evolved to address spectral variability and nonlinear effects (Bioucas-Dias et al., 2012). Traditional methods, such as compressed sensing (Donoho, 2006), least angle regression (Efron et al., 2004), basis pursuit and its denoising variants (Chen et al., 2001) and matching pursuit (Mallat and Zhang, 1993), utilized sparse regression frameworks to select relevant signatures from established libraries, assuming linear combinations of pure endmembers. Despite strong theoretical foundations, these approaches faced limitations in library acquisition and calibration (Iordache et al., 2011). This led to dictionary learning approaches (Aharon et al., 2006; Elad and Aharon, 2006) that extract spectral information directly from datasets, demonstrating capability to approximate nonlinear manifolds for individual materials without requiring explicit nonlinear models (Charles et al., 2011). While modern machine learning approaches offer adaptive, data-driven representations that capture material-specific variations (Bioucas-Dias et al., 2012), they primarily address local spectral variability rather than complex multiple scattering effects from three-dimensional structures. This limitation becomes evident when comparing with specialized methods like US-DART, which specifically accounts for multiple scattering phenomena—a persistent challenge for regression-based methods even with machine learning enhancements. The effectiveness of unmixing strategy depends on its alignment with the physical characteristics of the observed scene.

US-DART shows promising advantages over existing deep learning methods in terms of physical interpretability. Unlike neural network approaches (Li et al., 2024), the gradients obtained by US-DART have clear physical significance. Similarly, while physical-based Monte Carlo computer graphics models like Mitsuba 3 (Vicini et al., 2021), retrieve

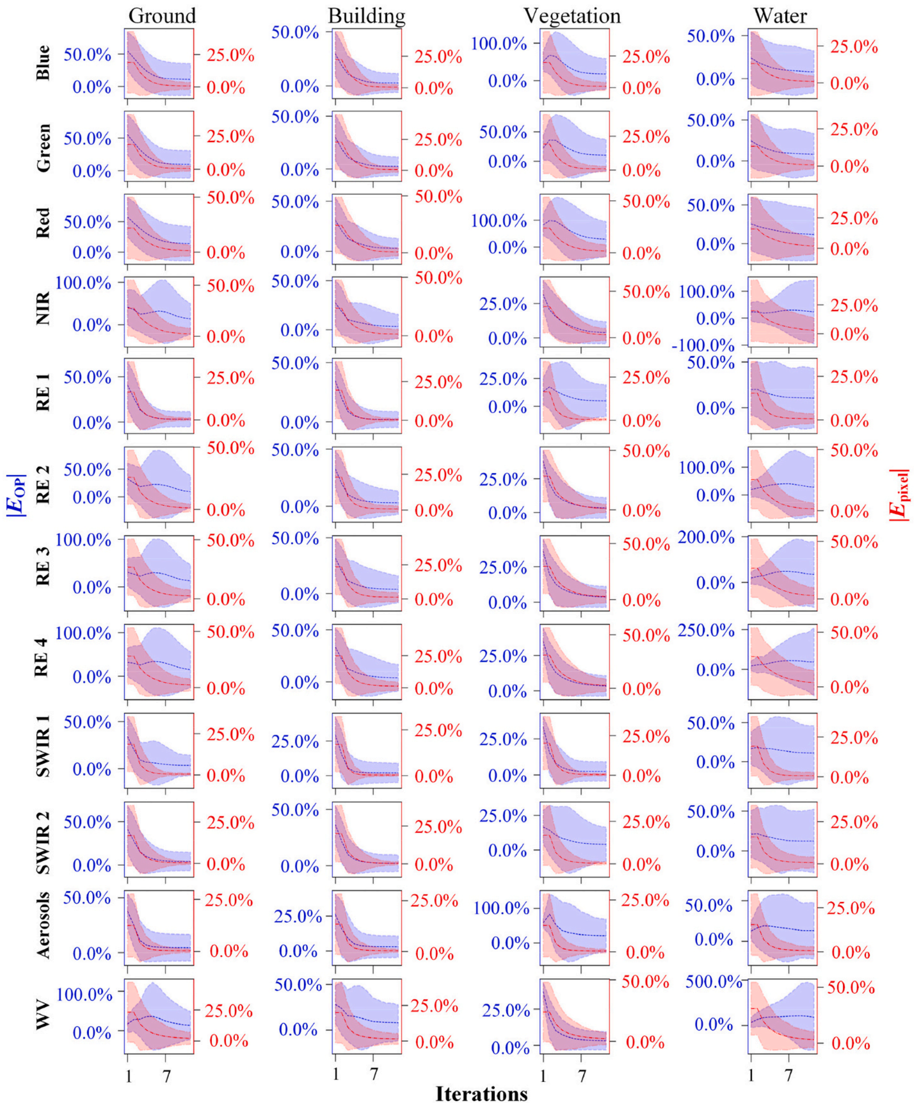


Fig. 25. The median error variations with iterations for the Basel scene. The left blue y-axis shows the absolute values of relative error of retrieved optical property maps. The right red y-axis shows the absolute values of the relative error of simulated reflectance images. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

material surface OP via backward ray tracing and gradient descent (Kingma and Ba, 2015), US-DART's multi-pixel method maintains better spatial continuity of OP and achieves convergence with fewer iterations, demonstrating its computational efficiency for complex scattering interactions.

4.1.2. Limitations and prospects

Like other spatial correlation-based unmixing methods, US-DART has inherent limitations from the multi-pixel unmixing methods. These algorithms necessitate numerous adjacent pixels with varying surface parameters. In entirely uniform scenes, the gradients from different pixels are nearly identical, resulting in a strong linear correlation within the equation system. Moreover, the application of autocorrelation of OP among neighboring pixels inevitably leads to a reduction in spatial resolution (Zhan et al., 2013). Moreover, the US-DART method is unable to discern differences in the OP of the same element in a pixel.

Like other multi-source data-based unmixing methods, US-DART has inherent limitations for using a 3D mock-up (Granero-Belinchon et al., 2020). This approach suffers from geometric misalignment between DART-simulated and satellite images and variations in the viewing angle. These factors can lead to errors in gradient estimation, affecting OP retrieval accuracy. Additionally, the requirement of the 3D mock-up increases the complexity of using US-DART. However, the emergence of LiDAR technology has made these 3D mock-ups more accessible. For example, Blender, a commercial software, offers a free plugin (https://blender-addons.org/blosm/) that enables users to download urban 3D mock-ups from OpenStreetMap worldwide. These mock-ups usually have a long usability span as they typically remain stable over several years.

US-DART is also inevitably subject to the limitations of errors from atmospheric correction and sensor noise, as it utilizes surface reflectance images. Future enhancements could involve the direct use of top-of-atmosphere reflectance images and integrating the atmosphere radiation transfer module (Wang and Gastellu-Etchegorry, 2021) into US-DART.

Since US-DART employs a 3D radiative transfer model to compute gradients, it necessitates ray tracing to determine radiation transfer for each facet in the urban scene. Given the vast number of facets involved, the computational load is substantial, resulting in lower efficiency compared with 2D unmixing. We are currently working on reducing the time of gradient calculation.

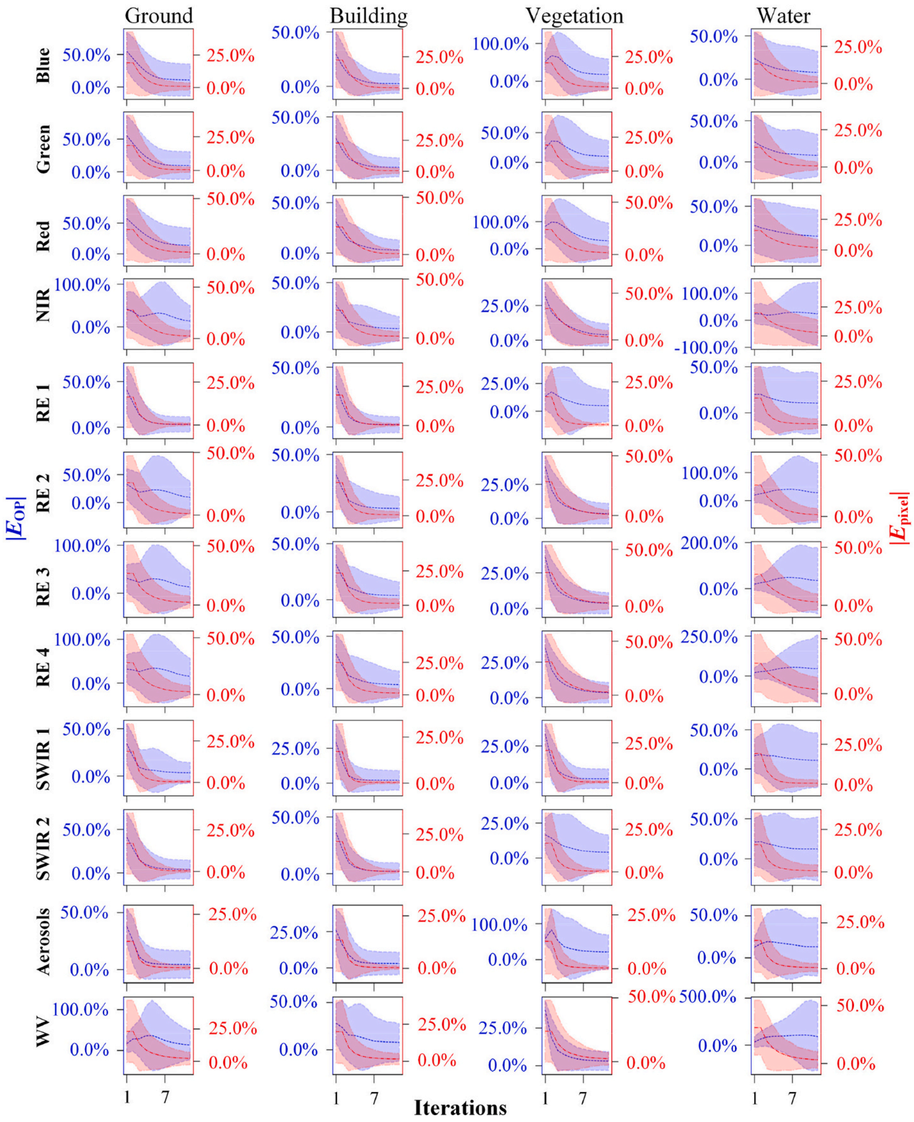


Fig. 26. The mean error variations with iterations for the Brussels scene. The left blue y-axis shows the absolute values of relative error of retrieved optical property maps. The right red y-axis shows the absolute values of relative error of simulated reflectance images. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.2. Conclusions

We propose the US-DART for retrieving element OP from mixed pixels in the satellite imagery. This method is assessed using two types of scenes (vegetation and urban) and image types (Sentinel-2 and pseudo-satellite images simulated by DART). Generally, the median relative error for reflectance is around 0.1 %, with opaque elements demonstrating higher accuracy than transparent ones. The accuracy of the reflectance image seems less influenced by the accuracy of the ratio between reflectance and transmittance of vegetation than that of the OP map. Assuming errors from co-registration and sensor noise are negligible, opaque elements typically exhibit an accuracy of around 1 % for OP error. However, the OP map accuracy of transparent elements heavily relies on the ratio's accuracy. If the ratio is accurate, the retrieved OP's accuracy can range between 1 % and 5 %.

US-DART, requiring a multispectral (or even mono-spectral) surface reflectance image, a 3D mock-up, and common DART parameters (e.g., spatial resolutions and skylight ratio) as input, eliminates the need for hyperspectral images and other restrictive conditions. The retrieved OP can be used as input parameters of DART for spatial resolution conversion, surface albedo simulation, and SW energy balance modeling (Letu et al., 2020; Wang et al., 2018). However, US-DART may encounter common issues like other multi-pixel and multi-source data models. Future enhancements to US-DART could involve directly using top-of-atmosphere reflectance images incorporating the atmosphere module in DART. US-DART, compatible with Windows and Unix platforms, features a concise graphical user interface. The source code of US-DART is distributed with the DART model and can be freely accessed at the website (https://dart.omp.eu/#/getDart) for academic and research purposes.

CRediT authorship contribution statement

Zhijun Zhen: Writing – original draft, Validation, Methodology, Funding acquisition, Data curation. **Shengbo Chen:** Writing – review & editing, Supervision. **Nicolas Lauret:** Software, Data curation. **Abdelaziz Kallel:** Methodology. **Eric Chavanon:** Software. **Tiangang Yin:** Writing – review & editing. **Jonathan León-Tavares:** Writing – review & editing, Funding acquisition. **Biao Cao:** Writing – review & editing. **Jordan Guilleux:** Software. **Jean-Philippe Gastellu-Etchegorry:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

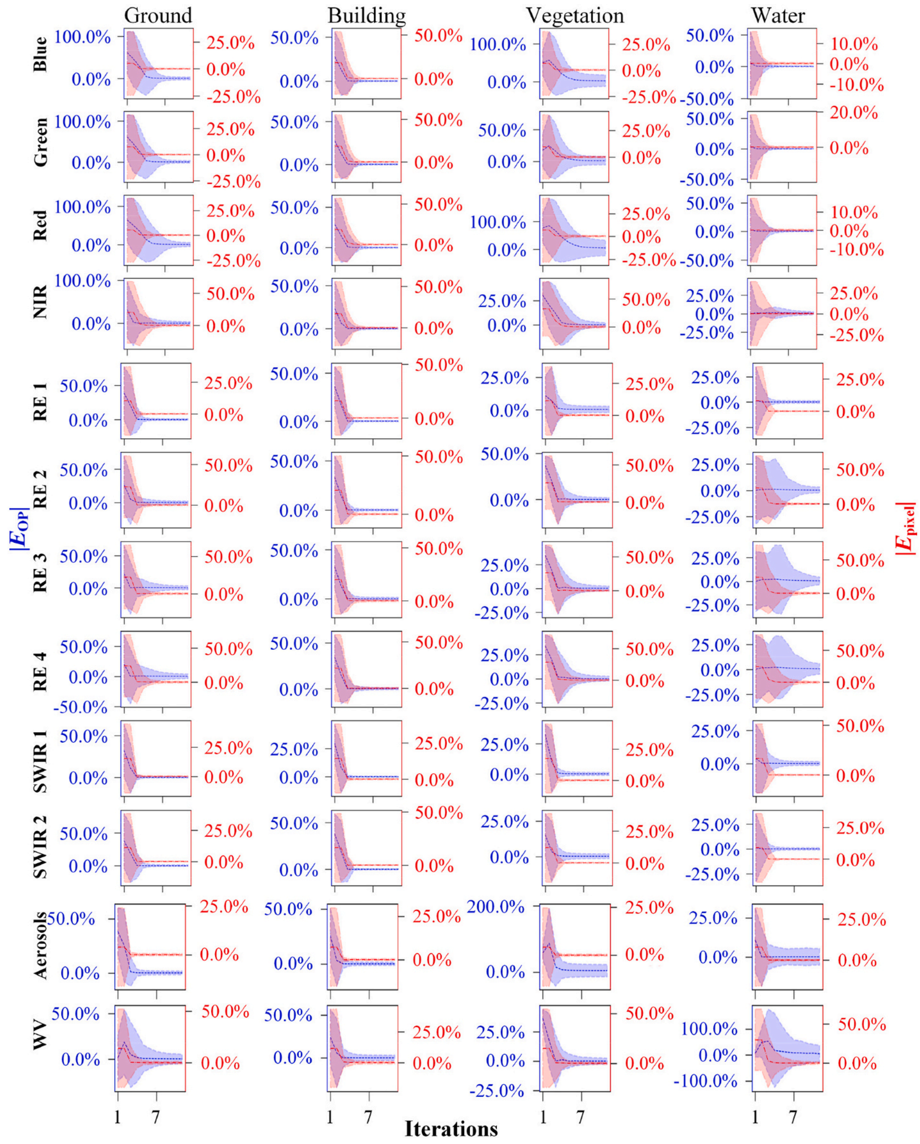


Fig. 27. The median error variations with iterations for the Brussels scene. The left blue y-axis shows the absolute values of relative error of retrieved optical property maps. The right red y-axis shows the absolute values of the relative error of simulated reflectance images. NIR: Near infrared. RE: Red edge. SWIR: Shortwave infrared. WV: Water vapor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Comparison of time consumption between US-DART and DART calibration for the vegetation scene before iteration 7. The following iterations show a similar trend (unit: second).

Iteration	Process	US-DART	DART calibration
Iteration 0	DART simulation	209.55	202.59
	Inversion	197.34	198.71
Iteration 1	DART simulation	205.80	206.88
	Inversion	7.13	7.39
Iteration 2	DART simulation	204.85	204.06
	Inversion	832.72	14.98
Iteration 3	DART simulation	215.74	202.64
	Inversion	9.96	17.49
Iteration 4	DART simulation	205.71	204.60
	Inversion	10.16	16.41
Iteration 5	DART simulation	209.68	202.61
	Inversion	9.90	13.56
Iteration 6	DART simulation	210.77	204.67
	Inversion	10.13	13.41
Iteration 7	DART simulation	213.85	204.61
	Inversion	10.02	13.38

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We understand that the corresponding author is the sole contact for the editorial process (including editorial manager and direct communications with the office). He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the corresponding author.

Acknowledgments

The authors thank the National Natural Science Foundation of China (NSFC, grant 42201372), the Belgium Science Policy Office BELSPO STEREO IV Programme (contract SR/67/408), and the TOSCA program of the French Space Center (CNES).

The authors thank Robbe Neyns from Vrije Universiteit Brussel for providing Brussels's tree geometry and location data. The authors thank Roland Vogt and Christian Feigenwinter from University of Basel for providing Basel's data. The authors thank Lucas Landier from CNES for his innovative ideas on the inversion of satellite images.

This research utilized data provided by the European Space Agency (ESA) through the Copernicus program (Sentinel-2) and Planet Labs Inc. (PlanetScope). We are grateful for the support and access to these valuable datasets. The authors express their gratitude to the editor and reviewers for their valuable feedback and thorough review of the manuscript.

Data availability

The source code of US-DART is distributed with the DART model which is available at DART website (https://dart.omp.eu/#/getDart). The source code of GeFolki is available at website (https://github.com/aplyer/gefolki). Sentinel-2 is available at European Space Agency (https://browser.dataspace.copernicus.eu/). PlanetScope is available at Planet Labs Inc. (https://www.planet.com/). The Brussels' mock-up is available at Urbis (https://datastore.brussels/web/urbis-download).

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34, 2274–2282.

Aharon, M., Elad, M., Bruckstein, A., 2006. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process. 54, 4311–4322.

Bay, S., Macfarlane, R., Wayment, T., Bearman, G., 2010. Multi-spectral imaging vs. monospectral infrared imaging. Z. Papyrol. Epigr, 211–217.

Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., Chanussot, J., 2012. Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches. IEEE J. Sel. Top. Appl. Earth Observ. 5, 354–379.

Borsoi, R.A., Imbiriba, T., Bermudez, J.C.M., 2019. Deep generative endmember modeling: an application to unsupervised spectral unmixing. IEEE Trans. Comput. Imaging 6, 374–384.

Borsoi, R.A., Imbiriba, T., Bermudez, J.C.M., Richard, C., Chanussot, J., Drumetz, L., Tourneret, J.-Y., Zare, A., Jutten, C., 2021. Spectral variability in hyperspectral data unmixing: a comprehensive review. IEEE Geosci. Remote Sens. Mag. 9, 223–270.

Borsoi, R.A., Erdo, D., Imbiriba, T., 2023. Learning interpretable deep disentangled neural networks for hyperspectral Unmixing. IEEE Trans. Comput. Imaging 9, 977–991.

Brezini, S.E., Deville, Y., Karoui, M.S., Benhalouche, F.Z., Ouamri, A., 2021. A penalization-based NMF approach for hyperspectral unmixing addressing spectral variability with an additively-tuned mixing model. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, pp. 3841–3844.

Brigot, G., Colin-Koeniguer, E., Plyer, A., Janez, F., 2016. Adaptation and evaluation of an optical flow method applied to coregistration of forest remote sensing images. IEEE J. Sel. Top. Appl. Earth Observ. 9, 2923–2939.

Charles, A.S., Olshausen, B.A., Rozell, C.J., 2011. Learning sparse codes for hyperspectral imagery. IEEE J. Select. Top. Sign. Process. 5, 963–978.

Chen, J.M., Liu, J., 2020. Evolution of evapotranspiration models using thermal and shortwave remote sensing data. Remote Sens. Environ. 237, 111594.

Chen, S.S., Donoho, D.L., Saunders, M.A., 2001. Atomic decomposition by basis pursuit. SIAM Rev. 43, 129–159.

Chen, J.M., Xu, M., Wang, R., Li, D., Liu, R., Ju, W., Cheng, T., 2022. Next step in vegetation remote sensing: Synergetic retrievals of canopy structural and leaf biochemical parameters. In: New Think. GISci. Springer, pp. 207–220.

Chrysoulakis, N., Grimmond, S., Feigenwinter, C., Lindberg, F., Gastellu-Etchegorry, J.-P., Marconcini, M., Mitraka, Z., Stagakis, S., Crawford, B., Olofson, F., 2018. Urban energy exchanges monitoring from space. Sci. Rep. 8, 1–8.

Cline, A.K., Moler, C.B., Stewart, G.W., Wilkinson, J.H., 1979. An estimate for the condition number of a matrix. SIAM J. Numer. Anal. 16, 368–375.

Cooley, T., Anderson, G.P., Felde, G.W., Hoke, M.L., Ratkowski, A.J., Chetwynd, J.H., Gardner, J.A., Adler-Golden, S.M., Matthew, M.W., Berk, A., 2002. FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation. In: IEEE Int. Geosci. Remote Sens. Symp. IEEE, pp. 1414–1418.

Cubero-Castan, M., Chanussot, J., Achard, V., Briottet, X., Shimoni, M., 2014. A physics-based unmixing method to estimate subpixel temperatures on mixed pixels. IEEE Trans. Geosci. Remote Sens. 53, 1894–1906.

Dong, T., Liu, J., Liu, J., He, L., Wang, R., Qian, B., McNairn, H., Powers, J., Shi, Y., Chen, J.M., 2023. Assessing the consistency of crop leaf area index derived from seasonal Sentinel-2 and Landsat 8 imagery over Manitoba, Canada. Agric. For. Meteorol. 332, 109357.

Donoho, D.L., 2006. Compressed sensing. IEEE Trans. Inf. Theory 52, 1289–1306.

Drumetz, L., Veganzones, M.-A., Henrot, S., Phlypo, R., Chanussot, J., Jutten, C., 2016. Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability. IEEE Trans. Image Process. 25, 3890–3905.

Drumetz, L., Chanussot, J., Jutten, C., 2019a. Spectral unmixing: a derivation of the extended linear mixing model from the hapke model. IEEE Geosci. Remote S 17, 1866–1870.

Drumetz, L., Chanussot, J., Jutten, C., 2019b. Variability of the endmembers in spectral unmixing. In: Data Handling in Science and Technology. Elsevier, pp. 167–203.

Du, X., Zare, A., Gader, P., Dranishnikov, D., 2014. Spatial and spectral unmixing using the beta compositional model. IEEE J. Sel. Top. Appl. Earth Observ. 7, 1994–2003.

Dunford, N., Schwartz, J.T., Bade, W.G., Bartle, R.G., 1963. Spectral Theory: Self Adjoint Operators in Hilbert Space. Interscience Publishers, New York.

Duthoit, S., Demarez, V., Gastellu-Etchegorry, J.-P., Martin, E., Roujean, J.-L., 2008. Assessing the effects of the clumping phenomenon on BRDF of a maize crop based on 3D numerical scenes using DART model. Agric. For. Meteorol. 148, 1341–1352.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least Angle Regression.

Elad, M., Aharon, M., 2006. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. 15, 3736–3745.

Feigenwinter, C., Parlow, E., Vogt, R., Schmutz, M., Chrysoulakis, N., Lindberg, F., Marconcini, M., Del Frate, F., 2017. Spatial distribution of sensible and latent heat flux in the URBANFLUXES case study city Basel (Switzerland). In: Joint Urban Remote Sens. Event. IEEE, pp. 1–4.

Feigenwinter, C., Vogt, R., Parlow, E., Lindberg, F., Marconcini, M., Del Frate, F., Chrysoulakis, N., 2018. Spatial distribution of sensible and latent heat flux in the city of Basel (Switzerland). IEEE J. Sel. Top. Appl. Earth Observ. 11, 2717–2723.

Fontanilles, G., Briottet, X., 2011. A nonlinear unmixing method in the infrared domain. Appl. Opt. 50, 3666–3677.

Franke, J., Roberts, D.A., Halligan, K., Menz, G., 2009. Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. Remote Sens. Environ. 113, 1712–1723.

Gastellu-Etchegorry, J.-P., 2008. 3D modeling of satellite spectral images, radiation budget and energy budget of urban landscapes. Meteorog. Atmos. Phys. 102, 187–207.

Gastellu-Etchegorry, J.-P., Yin, T., Lauret, N., Cajgfinger, T., Gregoire, T., Grau, E., Feret, J.-B., Lopes, M., Guilleux, J., Dedieu, G., 2015. Discrete anisotropic radiative transfer (DART 5) for modeling airborne and satellite spectroradiometer and LIDAR acquisitions of natural and urban landscapes. Remote Sens. 7, 1667–1701.

Granero-Belinchon, C., Michel, A., Achard, V., Briottet, X., 2020. Spectral unmixing for thermal infrared multi-spectral airborne imagery over urban environments: day and night synergy. Remote Sens. 12, 1871.

Guo, H., Ye, D., Xu, H., Bruzzone, L., 2024. OBSUM: an object-based spatial unmixing model for spatiotemporal fusion of remote sensing images. Remote Sens. Environ. 304, 114046.

Hapke, B., 2012. Theory of Reflectance and Emittance Spectroscopy. Cambridge University Press.

Hong, D., Yokoya, N., Chanussot, J., Zhu, X.X., 2018. An augmented linear mixing model to address spectral variability for hyperspectral unmixing. IEEE Trans. Image Process. 28, 1923–1938.

Houborg, R., McCabe, M.F., 2018. A cubesat enabled spatio-temporal enhancement method (cestem) utilizing planet, landsat and modis data. Remote Sens. Environ. 209, 211–226.

Imbiriba, T., Borsoi, R.A., Bermudez, J.C.M., 2018. Generalized linear mixing model accounting for endmember variability. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 1862–1866.

Iordache, M.-D., Bioucas-Dias, J.M., Plaza, A., 2011. Sparse unmixing of hyperspectral data. IEEE Trans. Geosci. Remote Sens. 49, 2014–2039.

Keshava, N., Mustard, J.F., 2002. Spectral unmixing. IEEE Signal. Proc. Mag. 19, 44–57.

Kingma, D.P., Ba, J., 2015. Adam: a method for stochastic optimization. In: Int. Conf. Learn. Represent.

Landier, L., 2018. Modélisation 3D du bilan radiatif des milieux urbains par inversion d'images satellites en cartes de réflectance et de température des matériaux urbains. In: Centre d'Etudes Spatiales de la Biosphère. Université de Toulouse, Toulouse.

Landier, L., Al Bitar, A., Lauret, N., Gastellu-Etchegorry, J.-P., Aubert, S., Mitraka, Z., Feigenwinter, C., Parlow, E., Heldens, W., Kotthaus, S., 2016. 3D modeling of radiative transfer and energy balance in urban canopies combined to remote sensing acquisitions. In: IEEE Int. Geosci. Remote Sens. Symp. IEEE, pp. 6738–6741.

Landier, L., Gastellu-Etchegorry, J., Al Bitar, A., Chavanon, E., Lauret, N., Feigenwinter, C., Mitraka, Z., Chrysoulakis, N., 2018. Calibration of urban canopies albedo and 3D shortwave radiative budget using remote-sensing data and the DART model. Eur. J. Remote Sens. 51, 739–753.

Leng, J., Chen, J.M., Li, W., Luo, X., Xu, M., Liu, J., Wang, R., Rogers, C., Li, B., Yan, Y., 2023. Global datasets of hourly carbon and water fluxes simulated using a satellite-based process model with dynamic parameterizations. Earth Syst. Sci. Data Discuss. 2023, 1–23.

Letu, H., Shi, J., Li, M., Wang, T., Shang, H., Lei, Y., Ji, D., Wen, J., Yang, K., Chen, L., 2020. A review of the estimation of downward surface shortwave radiation based on satellite data: methods, progress and problems. Sci. China Earth Sci. 63, 774–789.

Li, H., Borsoi, R.A., Imbiriba, T., Closas, P., Bermudez, J.C., Erdoğmuş, D., 2021. Model-based deep autoencoder networks for nonlinear hyperspectral unmixing. IEEE Geosci. Remote S 19, 1–5.

Li, Z., Yuan, Q., Yang, Q., Li, J., Zhao, T., 2024. Differentiable modeling for soil moisture retrieval by unifying deep neural networks and water cloud model. Remote Sens. Environ. 311, 114281.

Liu, Y., Chen, J.M., He, L., Wang, R., Smith, N.G., Keenan, T.F., Rogers, C., Li, W., Leng, J., 2023. Global photosynthetic capacity of C3 biomes retrieved from solar-induced chlorophyll fluorescence and leaf chlorophyll content. Remote Sens. Environ. 287, 113457.

Mahmoud, M.I., Dessouky, M.I., Deyab, S., Elfouly, F.H., 2007. Comparison between haar and daubechies wavelet transformions on FPGA technology. Int. J. Aerosp. Mech. Eng. 1, 141–145.

Mallat, S.G., Zhang, Z., 1993. Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Process. 41, 3397–3415.

Marta, S., 2020. Planet Imagery Product Specifications. In: Planet Labs: San Francisco, CA, USA. PLANET, Papua Province, Indonesia, p. 91.

Mitraka, Z., Chrysoulakis, N., Kamarianakis, Y., Partsinevelos, P., Tsouchlaraki, A., 2012. Improving the estimation of urban surface emissivity based on sub-pixel classification of high resolution satellite imagery. Remote Sens. Environ. 117, 125–134.

Myers, E., Kerekes, J., Daughtry, C., Russ, A., 2019. Assessing the impact of satellite revisit rate on estimation of corn Phenological transition timing through shape model fitting. Remote Sens. 11, 2558.

Nascimento, J.M., Dias, J.M., 2005. Vertex component analysis: a fast algorithm to unmix hyperspectral data. IEEE Trans. Geosci. Remote Sens. 43, 898–910.

Palsson, B., Sigurdsson, J., Sveinsson, J.R., Ulfarsson, M.O., 2018. Hyperspectral unmixing using a neural network autoencoder. IEEE Access 6, 25646–25656.

Plyer, A., Colin-Koeniguer, E., Weissgerber, F., 2015. A new coregistration algorithm for recent applications on urban SAR images. IEEE Geosci. Remote S 12, 2198–2202.

Qian, Y., Xiong, F., Qian, Q., Zhou, J., 2020. Spectral mixture model inspired network architectures for hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. 58, 7418–7434.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R., 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. Remote Sens. Environ. 65, 267–279.

Shen, P., Zhao, S., Ma, Y., Liu, S., 2023. Urbanization-induced Earth's surface energy alteration and warming: a global spatiotemporal analysis. Remote Sens. Environ. 284, 113361.

Shi, S., Zhao, M., Zhang, L., Altmann, Y., Chen, J., 2021. Probabilistic generative model for hyperspectral unmixing accounting for endmember variability. IEEE Trans. Geosci. Remote Sens. 60, 1–15.

Somers, B., Zortea, M., Plaza, A., Asner, G.P., 2012. Automated extraction of image-based endmember bundles for improved spectral unmixing. IEEE J. Sel. Top. Appl. Earth Observ. 5, 396–408.

Su, Y., Li, J., Plaza, A., Marinoni, A., Gamba, P., Chakravortty, S., 2019. DAEN: deep autoencoder networks for hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. 57, 4309–4321.

Thouvenin, P.-A., Dobigeon, N., Tourneret, J.-Y., 2015. Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. IEEE Trans. Signal Process. 64, 525–538.

Uezato, T., Murphy, R.J., Melkumyan, A., Chlingaryan, A., 2015. A novel spectral unmixing method incorporating spectral variability within endmember classes. IEEE Trans. Geosci. Remote Sens. 54, 2812–2831.

Uezato, T., Murphy, R.J., Melkumyan, A., Chlingaryan, A., 2016. Incorporating spatial information and endmember variability into unmixing analyses to improve abundance estimates. IEEE Trans. Image Process. 25, 5563–5575.

Uezato, T., Fauvel, M., Dobigeon, N., 2018. Hyperspectral image unmixing with LiDAR data-aided spatial regularization. IEEE Trans. Geosci. Remote Sens. 56, 4098–4108.

Uezato, T., Yokoya, N., He, W., 2020. Illumination invariant hyperspectral image unmixing based on a digital surface model. IEEE Trans. Image Process. 29, 3652–3664.

Vicini, D., Speierer, S., Jakob, W., 2021. Path replay backpropagation: differentiating light paths using constant memory and linear time. ACM Trans. Graph. 40, 1–14.

Wang, Y., Gastellu-Etchegorry, J.-P., 2021. Accurate and fast simulation of remote sensing images at top of atmosphere with DART-lux. Remote Sens. Environ. 256, 112311.

Wang, T., Shi, J., Yu, Y., Husi, L., Gao, B., Zhou, W., Ji, D., Zhao, T., Xiong, C., Chen, L., 2018. Cloudy-sky land surface longwave downward radiation (LWDR) estimation by integrating MODIS and AIRS/AMSU measurements. Remote Sens. Environ. 205, 100–111.

Wang, M., Zhao, M., Chen, J., Rahardja, S., 2019. Nonlinear unmixing of hyperspectral data via deep autoencoder networks. IEEE Geosci. Remote S 16, 1467–1471.

Wang, X., Chen, J.M., Ju, W., Zhang, Y., 2022. Seasonal variations in leaf maximum photosynthetic capacity and its dependence on climate factors across global FLUXNET sites. J. Geophys. Res. Biogeosci. 127, e2021JG006709.

Wang, J., Chen, J.M., Qiu, F., Fan, W., Xu, M., Wang, R., 2024a. Simultaneous estimation of leaf directional-hemispherical reflectance and transmittance from multi-angular canopy reflectance. Remote Sens. Environ. 304, 114025.

Wang, Y., Kallel, A., Zhen, Z., Lauret, N., Guilleux, J., Chavanon, E., Gastellu-Etchegorry, J.-P., 2024b. 3D Monte Carlo differentiable radiative transfer with DART. Remote Sens. Environ. 308, 114201.

Welch, R., 1982. Spatial resolution requirements for urban studies. Int. J. Remote Sens. 3, 139–146.

Xie, X., Chen, J.M., Gong, P., Li, A., 2021. Spatial scaling of gross primary productivity over sixteen mountainous watersheds using vegetation heterogeneity and surface topography. J. Geophys. Res. Biogeosci. 126, e2020JG005848.

Xie, X., Li, A., Chen, J.M., Guan, X., Leng, J., 2022. Quantifying scaling effect on gross primary productivity estimation in the upscaling process of surface heterogeneity. J. Geophys. Res. Biogeosci. 127, e2021JG006775.

Xu, M., Shang, R., Chen, J.M., Zeng, L., 2023. LACC2. 0: improving the LACC algorithm for reconstructing satellite-derived time series of vegetation biochemical parameters. Remote Sens. 15, 3277.

Yin, T., Gastellu-Etchegorry, J.-P., Lauret, N., Grau, E., Rubio, J., 2013. A new approach of direction discretization and oversampling for 3D anisotropic radiative transfer modeling. Remote Sens. Environ. 135, 213–223.

Zhan, W., Chen, Y., Zhou, J., Wang, J., Liu, W., Voogt, J., Zhu, X., Quan, J., Li, J., 2013. Disaggregation of remotely sensed land surface temperature: literature survey, taxonomy, issues, and caveats. Remote Sens. Environ. 131, 119–139.

Zhang, J., Rivard, B., Sánchez-Azofeifa, A., Castro-Esau, K., 2006. Intra-and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: implications for species identification using HYDICE imagery. Remote Sens. Environ. 105, 129–141.

Zhang, G., Scheunders, P., Cerra, D., 2023. Shadow-aware nonlinear spectral unmixing with spatial regularization. IEEE Trans. Geosci. Remote Sens. 61, 5517516.

Zhen, Z., Gastellu-Etchegorry, J.-P., Chen, S., Yin, T., Chavanon, E., Lauret, N., Guilleux, J., 2021. Quantitative analysis of DART calibration accuracy for retrieving spectral signatures over urban area. IEEE J. Sel. Top. Appl. Earth Observ. 14, 10057–10068.

Zhou, C., Rodrigues, M.R., 2021. ADMM-based hyperspectral unmixing networks for abundance and endmember estimation. IEEE Trans. Geosci. Remote Sens. 60, 1–18.

Zhu, J., Shi, J., 2018. An algorithm for subpixel snow mapping: Extraction of a fractional snow-covered area based on ten-day composited avhrr\/2 data of the qinghai-tibet plateau. IEEE Geosci. Remote Sens. Mag. 6, 86–98.

Zhu, J., Shi, J., Wang, Y., 2012. Subpixel snow mapping of the Qinghai–Tibet Plateau using MODIS data. Int. J. Appl. Earth Obs. Geoinf. 18, 251–262.