






Article

Retrieval of Vegetation Indices and Vegetation Fraction in Highly Compact Urban Areas: A 3D Radiative Transfer Approach

Wenya Xue ¹, Liping Feng ¹, Jinxin Yang ^{1,*}, Yong Xu ¹, Hung Chak Ho ², Renbo Luo ¹, Massimo Menenti ^{3,4} and Man Sing Wong ^{5,6}

¹ School of Geography and Remote Sensing, Guangzhou University, Guangzhou 510006, China; 32101800037@e.gzhu.edu.cn (W.X.); 32101800015@e.gzhu.edu.cn (L.F.); xu1129@gzhu.edu.cn (Y.X.); luorb@gzhu.edu.cn (R.L.)

² Department of Public and International Affairs, City University of Hong Kong, Hong Kong 999077, China; hungcho2@cityu.edu.hk

³ Faculty of Civil Engineering and Earth Sciences, Delft University of Technology, 2600 GA Delft, The Netherlands; m.menenti@tudelft.nl

⁴ State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

⁵ Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong 999077, China; ls.charles@polyu.edu.hk

⁶ Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hong Kong 999077, China

* Correspondence: yangjx11@gzhu.edu.cn

Abstract: Vegetation indices, especially the normalized difference vegetation index (NDVI), are widely used in urban vegetation assessments. However, estimating the vegetation abundance in urban scenes using the NDVI has constraints due to the complex spectral signature related to the urban structure, materials and other factors compared to natural ground surfaces. This paper employs the 3D discrete anisotropic radiative transfer (DART) model to simulate the spectro-directional reflectance of synthetic urban scenes with various urban geometries and building materials using a flux-tracking method under shaded and sunlit conditions. The NDVI is calculated using the spectral radiance in the red (0.6545 μm) and near-infrared bands (0.865 μm). The effects of the urban material heterogeneity and 3D structure on the NDVI, and the performance of three NDVI-based fractional vegetation cover (FVC) inversion algorithms, are evaluated. The results show that the effects of the building material heterogeneity on the NDVI are negligible under sunlit conditions but not negligible under shaded conditions. The NDVI value of building components within synthetic scenes is approximately zero. The shaded road exhibits a higher NDVI value in comparison to the illuminated road because of scattering from adjacent pixels. In order to correct the effects of scattering caused by building geometry, the reflectance of the Landsat 8/OLI image is corrected using the sky view factor (SVF) and then used to calculate the FVC. Jilin-1 satellite images with high spatial resolution (0.5 m) are used to extract the vegetation cover and then aggregated to 30 m spatial resolution to calculate the FVC for validation. The results show that the RMSE is up to 0.050 after correction, while the RMSE is 0.169 before correction. This study makes a contribution to the understanding of the effects of the urban 3D structure and material reflectance on the NDVI and provides insights into the retrieval of the FVC in different urban scenes.

Keywords: NDVI; FVC; urban geometry; building materials; DART



Academic Editor: Ran Meng

Received: 19 October 2024

Revised: 6 December 2024

Accepted: 30 December 2024

Published: 3 January 2025

Citation: Xue, W.; Feng, L.; Yang, J.; Xu, Y.; Ho, H.C.; Luo, R.; Menenti, M.; Wong, M.S. Retrieval of Vegetation Indices and Vegetation Fraction in Highly Compact Urban Areas: A 3D Radiative Transfer Approach. *Remote Sens.* **2025**, *17*, 143. <https://doi.org/10.3390/rs17010143>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Monitoring urban green space is important since the existence of green space in urban areas provides multiple biophysical, aesthetic, environment and socioeconomic benefits [1]. For pure pixels, the identification of urban vegetation is based on land use classification using different spatial resolution remote-sensing and GIS data [2,3]. Urban areas are highly heterogeneous, however, and the radiometric data provided by space-borne sensors capture mixed pixels in most cases. The combination of the 3D structure of urban space with the spectral heterogeneity of urban objects leads to the complex spectro-directional signatures of the observed targets, even when observing pure pixels. The retrieval of a pixel signature must account for multiple scattering by surrounding urban facets.

For mixed pixels, spectral unmixing is a useful method, but the spectral signatures of pure component objects must be known [4–6]. Additionally, using vegetation indices to estimate the vegetation fractional abundance has been widely investigated in remote-sensing research, especially the use of the NDVI [7,8]. The vegetation fraction is generally written as a linear or non-linear function of the NDVI [7]. The NDVI is widely used in urban climate and environment research as an indicator of urban vegetation abundance for, e.g., emissivity estimation [9–11], or as an indicator of vegetation cooling [12].

Theoretically, an ideal vegetation index should only be sensitive to vegetative cover and insensitive to other factors, e.g., background spectral radiance, illumination and view directions, and surface geometry [13,14]. The NDVI is sensitive to the background spectral radiance when used to characterize mixed surface targets [14]. Urban surface properties vary and the spectral signatures are very different from those of natural surfaces [15]. Additionally, urban areas have very complex geometry. The interaction between buildings and vegetation also changes the spectro-directional signal observed by remote sensing, with an impact on the retrieved vegetation indices in urban areas.

The accurate quantification of urban vegetation cover is a fundamental component of urban ecosystem assessment and sustainable planning. Previous studies have concentrated on FVC estimation for land use types such as forests, agricultural lands and grasslands [16–18], and sometimes, semi-arid areas and heterogeneous ecological systems after fire were also explored [19,20]. However, there is a paucity of knowledge and methodological development on urban FVC estimation. Air-borne remote-sensing data are useful [21], and their finer resolution facilitates the identification of urban vegetation. However, their global application is also hindered by limitations in spatial and temporal coverage, high costs, and processing complexities. Recently, the development of FVC estimation methods based on space-borne remote-sensing data predominantly employs machine learning and multi-source data fusion. Song et al. [22] proposed a generic FVC estimation process that is capable of jointly utilizing multiple satellite data. This process improves the spatio-temporal resolution of FVC products. Jia et al. [23] employed radiative transfer models to simulate the GF-1 WFV canopy reflectance and FVC values and trained backpropagation neural networks with these synthetic samples. The algorithm was validated as being effective ($R^2 = 0.790$, $RMSE = 0.073$), but these methods are generic and have not been developed specifically for urban areas. Urban-object-based mixture analysis was performed by Cai et al. [24], which differs from traditional pixel-based mixture analysis, resulting in a higher urban FVC estimation accuracy ($R^2 = 0.92$, $RMSE = 0.0956$). Several studies used the NDVI to capture the urban vegetation cover condition [25,26], but the complex urban geometry in compact urban areas affects the solar radiation and radiative transfer, i.e., it modifies the inherent spectro-directional signature of the observed targets, including the red and near-infrared reflectance applied to obtain the NDVI. However, previous studies have not elucidated the manner in which the urban 3D geometry and building materials affect the NDVI of mixed pixels. The applicability of multiple NDVI-based FVC

estimation algorithms for diverse urban contexts has yet to be assessed, which is crucial to improve the urban FVC estimation accuracy.

This study is based on numerical experiments with the discrete anisotropic radiative transfer (DART) model to construct urban simulation scenes with varying geometrical and material features, which allows us to explore the relationship between the 3D geometry and material reflectance and the NDVI of mixed pixels. The DART model is one of the most comprehensive physically based 3D models of Earth–atmosphere radiative transfer, covering the spectral domain from ultraviolet to thermal infrared wavelengths. It simulates the optical 3D radiative budget (RB) and can be applied to calculate the optical signals captured by proximal, aerial, and satellite imaging spectroradiometers and laser scanners for any urban and/or natural landscapes and for any experimental and instrumental configurations [27]. The DART model has been validated through a series of RAMI (Radiation transfer Model Intercomparison) experiments [28–32], and its forward simulation of vegetation reflectance has also been rigorously verified by real measurements [33]. Zhen et al. [34] have conducted an extensive evaluation of the DART model’s calibration accuracy for the iterative inversion of shortwave satellite images, elucidating its potential as an instrumental tool for obtaining spectral signature maps of urban materials. This means the DART model can handle the shortwave radiative transfer in urban areas. The red and near-infrared bands employed to calculate the key variable (i.e., NDVI) in our study are encompassed within the shortwave spectral domain. Likewise, the utility of the DART model has been substantiated across scientific realms, and its iterative enhancements and extensive applications have been demonstrated in detail by Gastellu-Etchegorry et al. [35]. These include the development of remote-sensing inversion techniques and the design of satellite sensors and novel vegetation indices. Yan et al. [36] used DART to model the longwave radiation in a complex-relief terrain. Zhang et al. [16] used DART to model the urban radiative transfer to retrieve the effective emissivity. Dissegna et al. [37] devised an innovative approach to produce detailed modeling of the mean radiant temperature (T_{mrt}) spatial distribution using the DART model. The effectiveness of this methodology was substantiated through a case study in Singapore, a tropical city, where the DART-modeled T_{mrt} was compared with field-estimated values. This means the DART model can handle the radiative transfer in any terrain; therefore, it is a promising candidate for developing and enhancing inversion methods in urban areas with vegetation and buildings. In this study, the research questions are as follows. (1) How does the urban geometry and materials affect the NDVI in urban areas? (2) How should the NDVI be applied to estimate the vegetation fractional abundance in urban areas?

Section 2 describes the model used in this study, the simulation setup for the scenes, and the formulas used. Section 3 presents a compilation and analysis of the experimental results, assessing the effects of different factors on the NDVI and FVC, as well as the model validation. Comments on the results are articulated in Section 4 and conclusions will be presented in Section 5.

2. Materials and Methods

2.1. DART Model

Since its development in 1992, the DART model (<https://dart.omp.eu>, accessed on 15 October 2024) has been one of the most comprehensive and accurate radiative transfer models for simulating the 3D radiation budget [38]. The DART model comprises three principal modules: DART-FT, DART-RC, and DART-Lux [35]. The DART-FT module was selected for our study, which uses a flux tracking (FT) method to track radiation in a finite number of directions in a voxelized scene (i.e., discrete coordinate method) [8]. The 3D scene elements in this module are comprised of a voxel array (e.g., vegetation, atmosphere)

and facets (e.g., roads, roofs, walls). These elements possess specific optical properties (i.e., reflectance and transmittance), and an SQL database of the spectral properties of elements exists in the DART model.

The graphical user interface (GUI) enables users to define the input parameters (e.g., atmospheric properties, scene properties, sun position, sensor view direction) and anticipated output results (e.g., images, radiative budget, lidar waveform). Subsequently, remote-sensing indices (e.g., NDVI) can be calculated based on the output results.

2.1.1. Input Parameters

The setting of the DART parameters in our numerical experiments is presented in Table 1. To simulate the radiative transfer from the top of the atmosphere to the bottom of the atmosphere, an irradiance spectral database called “Solar_constant” is employed to characterize the atmospheric properties. We determine the “sun position” by setting the temporal parameters, and two distinct types of sun angle were considered for all the scenes: (1) scenes with a shadow condition: the temporal parameter was set at 10:00 am on the 22nd of June (local solar time), while the solar elevation angle was 56.07° (Figure 1b); and (2) scenes under a sunlit condition: the temporal parameter was set at 12:00 pm on the same date, while the solar elevation angle was 83.5° (Figure 1c).

Table 1. The DART model’s input parameters.

| Module | Value |
|-----------------------|---|
| Atmosphere Property | Irradiance Spectral Database Irradiance Table Irradiance Model Solar_constant TOASolar_THKUR irradiance |
| Scene Property | Location Size Latitude: 23°25′48″N, Longitude: 113°15′36″E Scene size: 40 m × 50 m, Cell size: 0.5 m × 0.5 m × 0.5 m |
| Sun Position | Shadow Condition Sunlit Condition Elevation: 56.07°, Azimuth: 82.41° Elevation: 83.5°, Azimuth: 88.67° |
| Sensor View Direction | View Angle Elevation: 90° |
| Spectral Band | Red Band Central wavelength: 0.6545 μm Spectral bandwidth: 0.037 μm Near-infrared Band Central wavelength: 0.865 μm Spectral bandwidth: 0.028 μm |

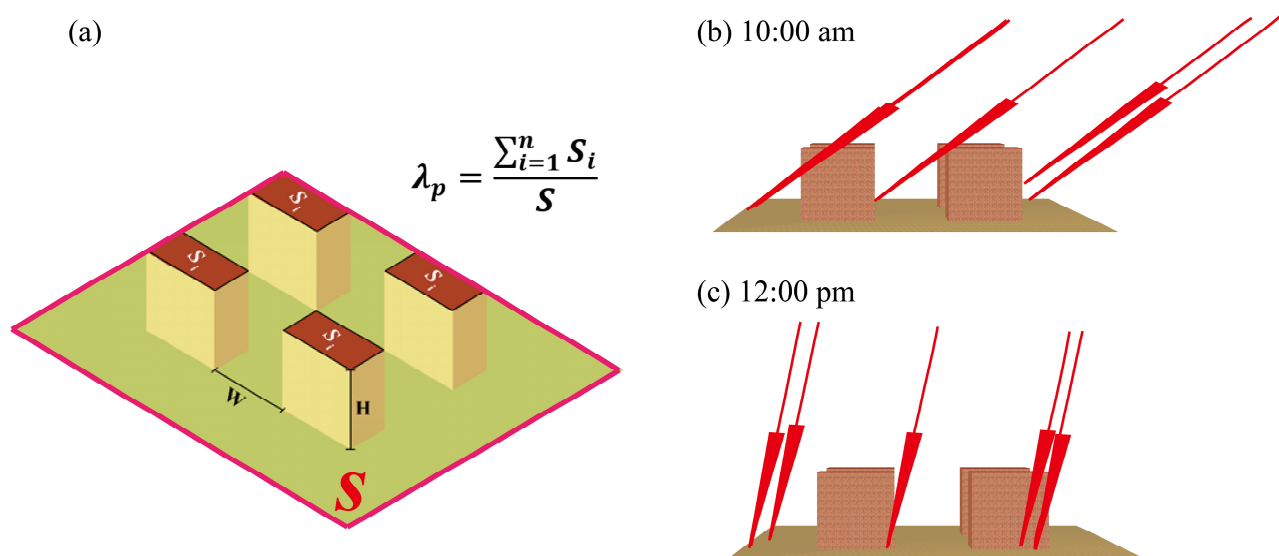


Figure 1. Scene schematic: (a) urban scene with building height (H), street width (W), roof area (S_i) and total horizontal area (S); (b) urban scene with sunlight under shadow conditions; and (c) urban scene with sunlight under sunlit conditions.

Our aim is to construct a remote-sensing observation model that can be practically applied, and the sensor view direction of the Landsat series satellites is nearly vertical. Thus, only data from the nadir view direction will be analyzed in this study. All the facets and voxel array reflectance properties are assumed to be Lambertian (Table 1).

The reflectance in the red and near-infrared bands was simulated using the DART model for three types of scenes: mixed scenes, vegetation-only scenes, and building-only scenes (Figure 2). The central wavelengths and widths of the red and near-infrared bands are referenced to Landsat 8/OLI (Table 1). Each scene has a constant parameter setting. The road material is designated as “loam_sandy_brown_bright_clean_smooth_arizona” for both the mixed and building-only scenes, and “sandy_loam_brown” for the vegetation-only scenes. A single building orientation and the DART model’s “repeated” mode are set to generate an infinite repetition of the actual cityscape model, thus avoiding unnecessary complications of urban scenes. The buildings in the actual scene are simplified into four regular rectangles of equal size, and the vegetation is simplified into a uniform type and morphology, with all the buildings and vegetation distributed evenly and symmetrically throughout the scene.

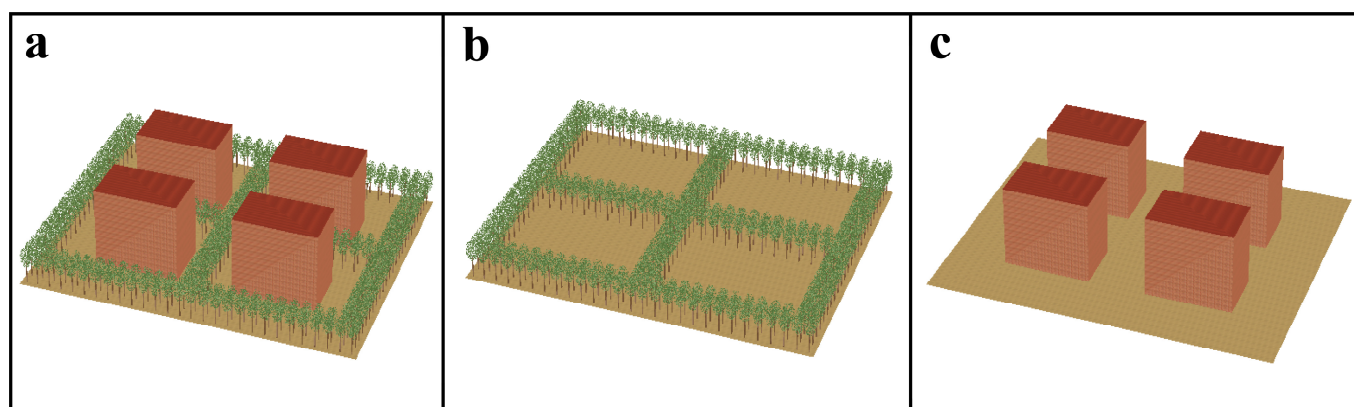


Figure 2. Three types of scenes created in the DART model are shown schematically: (a) mixed scene; (b) vegetation-only scene; and (c) building-only scene.

Zhang et al. [39] analyzed the reflectance of over 30 building materials from the ASTER Spectral Library (<https://speclib.jpl.nasa.gov>, accessed on 15 October 2024) and found that material reflectance in the range from 0.1 to 0.5 is closer to reality. We sampled 0.3, 0.6, 0.9, 1.0, 1.2, and 1.5 times the material reflectance (R_{stdW}) of “brick_cement_yellow_new” in the DART model’s SQL database, thus encompassing a range of material reflectance between approximately 0.1 and 0.5 (Table 2). To study the impact of building materials on the NDVI and to take into account urban space’s greater variability, two types of buildings are identified:

- (1) Heterogeneous buildings ($R_{wall} \neq R_{roof}$): walls and roofs are made of different materials, the roof material is always designated as “roof_tile_ceramic_red_new”, while the wall facet reflectance (R_{wall}) is set to 0.3, 0.6, 0.9, 1.0, 1.2 or 1.5 times the R_{stdW} , respectively.
- (2) Homogeneous buildings ($R_{wall} = R_{roof}$): the materials of the wall and roof are always the same, and the material reflectance is 0.3, 0.6, 0.9, 1.0, 1.2 or 1.5 times the R_{stdW} , respectively.

Table 2. Reflectance in the red and near-infrared bands corresponds to different R_{wall} s.

| Band | 0.3 R_{stdW} | 0.6 R_{stdW} | 0.9 R_{stdW} | 1.0 R_{stdW} | 1.2 R_{stdW} | 1.5 R_{stdW} |
|---|----------------|----------------|----------------|----------------|----------------|----------------|
| Red (0.636 μm –0.673 μm) | 0.0999 | 0.1999 | 0.2998 | 0.3331 | 0.3997 | 0.4997 |
| NIR (0.851 μm –0.879 μm) | 0.1000 | 0.2000 | 0.3001 | 0.3334 | 0.4001 | 0.5001 |

To study the impact of urban geometry on the NDVI, two variables were selected: (1) the building height to street width ratio (H/W); and (2) the roof area index (λ_p) (Figure 1a):

$$\lambda_p = \frac{\sum_{i=1}^n S_i}{S} \quad (1)$$

The numerator represents the sum of the roof areas and the denominator is the scene's total horizontal area.

Stewart and Oke et al. [40] established the range of values for each indicator across diverse local climate zone categories through rigorous validation in several cities. Our geometric parameter values encompass the range of indicators for the majority of urban local climate zones. A total of 25 geometric 3D simulation scenes are constructed with H/W of 0.75, 1, 1.5, 3, and 4, and λ_p of 0.1, 0.1875, 0.3, 0.4375, and 0.6, respectively (Figure A1).

The vegetation parameters are presented in Table 3. Given that a high building density also constrains the urban greening area, with the area of 40 m \times 50 m for all the scenes (Table 1) and the maximum value of λ_p being 0.6 (Table 4), the upper limit of the FVC was set at 0.3, divided into equal segments by 0.05.

Table 3. The DART model's vegetation parameters.

| Module | Scenes With Vegetation |
|-----------------|---|
| FVC | 0.1/0.15/0.2/0.25/0.3 |
| Leaf Area Index | 0.30/0.45/0.60/0.75/0.90 |
| Vegetation Type | Leaf type: grass_rye Trunk type: bark_deciduous |
| Tree | Height below crown: 2.5 m Diameter below crown: 0.166 m Height within the tree crown: 1.5 m |
| Tree Crown | Crown shape: ellipsoid Crown height: 3.5 m First axis: 1.663 m Second axis: 1.663 m |

Table 4. The DART model's building parameters.

| Parameters | Scenes With Building |
|-------------------|--|
| Building Material | Heterogeneous building ($R_{wall} \neq R_{roof}$) $R_{roof}, R_{wall} = 0.3/0.6/0.9/1.0/1.2/1.5 R_{stdW}$ Homogeneous building ($R_{wall} = R_{roof}$) $R_{wall} = R_{roof} = 0.3/0.6/0.9/1.0/1.2/1.5 R_{stdW}$ |
| Building Geometry | H/W 0.75/1.0/1.5/3.0/4.0 λ_p 0.1/0.1875/0.3/0.4375/0.6 |

2.1.2. Output Parameters

The DART model generates reflectance in the red and near-infrared bands. The NDVI is calculated using Equation (2):

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (2)$$

ρ_{NIR} , ρ_{Red} are the nadir reflectance in the near-infrared and red bands, respectively.

2.2. Retrieval of FVC

This study evaluated the impact of the urban structure on FVC estimates using the NDVI and selected three methods for this purpose.

A classical FVC calculation method was proposed by Gutman and Ignatov et al. [41] (Equation (3)):

$$FVC_{-1} = \frac{NDVI_{mean} - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \quad (3)$$

In this study, we focused on calculating the FVCs for mixed scenes, with each scene comprising a specific number of cells (Table 1). To calculate the FVC for a specific mixed scene, $NDVI_{mean}$ represents the mean NDVI for all the cells within the mixed scene, $NDVI_{max}$ represents the maximum NDVI for a vegetation-only scene cell, while $NDVI_{min}$ represents the minimum NDVI for a building-only scene cell. Figure 2 shows a scene schematic of the three types of scenes (i.e., mixed scene, vegetation-only scene, and building-only scene).

Additionally, two semi-empirical relationships proposed by Frederic et al. [42] (Equation (4)) and Carlson and Ripley et al. [43] (Equation (5)) were applied:

$$FVC_{-2} = 1 - \left(\frac{NDVI_{max} - NDVI_{mean}}{NDVI_{max} - NDVI_{min}} \right)^{0.6175} \quad (4)$$

$$FVC_{-3} = \left(\frac{NDVI_{mean} - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right)^2 \quad (5)$$

The accuracy of FVC estimates using these three methods was evaluated using the difference (dif_{FVC}) between the FVC estimates and the FVC_{true} defined by the scene design (Table 2):

$$dif_{FVC} = FVC - FVC_{true} \quad (6)$$

FVC is calculated using Equations (3)–(5). The derived dif_{FVC} s (dif_{FVC-1} , dif_{FVC-2} , dif_{FVC-3}) are represented by the absolute values of the corresponding dif_{FVC} s, which are designated as $|dif_{FVC-1}|$, $|dif_{FVC-2}|$, and $|dif_{FVC-3}|$ (Section 3.3).

The root mean square error (RMSE) and bias (BIAS) were also used to assess the FVC estimation accuracy.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (FVC_{true} - FVC_i)^2} \quad (7)$$

$$BIAS = \frac{1}{N} \sum_{i=1}^N (FVC_i - FVC_{true}) \quad (8)$$

2.3. Correction Model for NDVI in Urban Areas

Taking into account the multiple scattering within urban pixels, a new method was constructed to calculate the urban vegetation cover. According to [44], on page 56, the sky view factor (SVF) can be employed to characterize the urban geometry, thereby facilitating the parameterization of the impact of urban geometry on radiative transfer. The reflectance of pixel(i) can be written as follows:

$$\alpha(i) = \frac{\alpha(i)' * SVF(i)}{1 - \alpha(i)' * [1 - SVF(i)]} \quad (9)$$

$\alpha(i)$ represents the reflectance in the red and near-infrared bands from satellite data. $\alpha(i)'$ denotes the corrected reflectance in the aforementioned bands and is calculated by

Equation (9). $SVF(i)$ is the scene-averaged SVF calculated from the urban digital surface model (DSM) using a toolbox [45].

2.4. Model Validation

We chose Hong Kong's Kowloon Peninsula as the study area to validate the estimated FVC based on the corrected reflectance using Equation (9). The effectiveness of the correction was evaluated by introducing the “true FVC” and comparing the RMSE (Equation (7)) and BIAS (Equation (8)) of the “original FVC” and “modified FVC” (Section 3.4). The three key FVCs were calculated as follows:

- (1) True FVC: the L3C class image product with a spatial resolution of 0.78 m from Jilin-1 (<https://www.jl1mall.com/>, accessed on 15 October 2024) on 10 May 2020 was selected for analysis. The FVC values were calculated and aggregated to 30 m.
- (2) Original FVC: the Level2 Tier1 class image product with a spatial resolution of 30 m from Landsat 8 (<https://earthexplorer.usgs.gov>, accessed on 15 October 2024) on 17 May 2020 was used as the source data for calculating the original FVC values.
- (3) Modified FVC: The SVF was calculated using DSM datasets (https://portal.crsdi.gov.hk/geoportal/?lang=en&datasetId=cedd_rcd_1629267205233_87895, accessed on 15 October 2024) with a spatial resolution of 0.5 m in 2020 and resampled to 30 m. The reflectance in the red and near-infrared bands of the Landsat 8 image was corrected using Equation (9), after which the modified FVC values could be calculated.

3. Results

To study the effects of urban geometry on the NDVI, the $NDVI_{mean}$ of the vegetation-only scenes with different FVCs was used as a reference (Table 5).

Table 5. $NDVI_{mean}$ of vegetation-only scenes with different FVCs.

| FVC | $NDVI_{mean}$ |
|------|---------------|
| 0.10 | 0.301 |
| 0.15 | 0.352 |
| 0.20 | 0.398 |
| 0.25 | 0.441 |
| 0.30 | 0.477 |

3.1. Effects of Urban Geometry on NDVI

To investigate the effects of H/W and λ_p on the NDVI, we computed the mean values of the $NDVI_{mean}$ for scenes with different H/W and λ_p but with the same material reflectance. Given the similar NDVI patterns observed in urban scenes with varying FVCs, only the $NDVI_{mean}$ with FVCs of 0.1 and 0.3 are presented in Figure 3.

It is evident that buildings cause the $NDVI_{mean}$ to be underestimated compared to the $NDVI_{mean}$ for vegetation-only scenes (Figure 3). All the curved lines have a lower $NDVI_{mean}$ than the one corresponding to the $NDVI_{mean}$ with the same FVC in Table 5. Figure 3 illustrates that shadows intensify the impact of urban geometry on the NDVI, with greater vegetation cover exhibiting a heightened degree of influence from urban geometry. A significant negative correlation is observed between the λ_p and the NDVI, regardless of the sunlight condition. However, the relationship between the H/W and the NDVI is found to be significantly affected by the sunlight conditions. Under sunlit conditions, the H/W and NDVI are almost independent, whereas with shadow conditions, the effects of the H/W on the NDVI are also related to the λ_p and FVC. When $\lambda_p = 0.1$ and 0.1875, the impact of the H/W on the NDVI was minimal. When $\lambda_p = 0.3, 0.4375$, and 0.6, there is a negative correlation between the H/W and the NDVI, with this correlation increasing in

line with the FVC. Thus, the NDVI may be underestimated to a greater extent in urban areas with high vegetation cover and high building density.

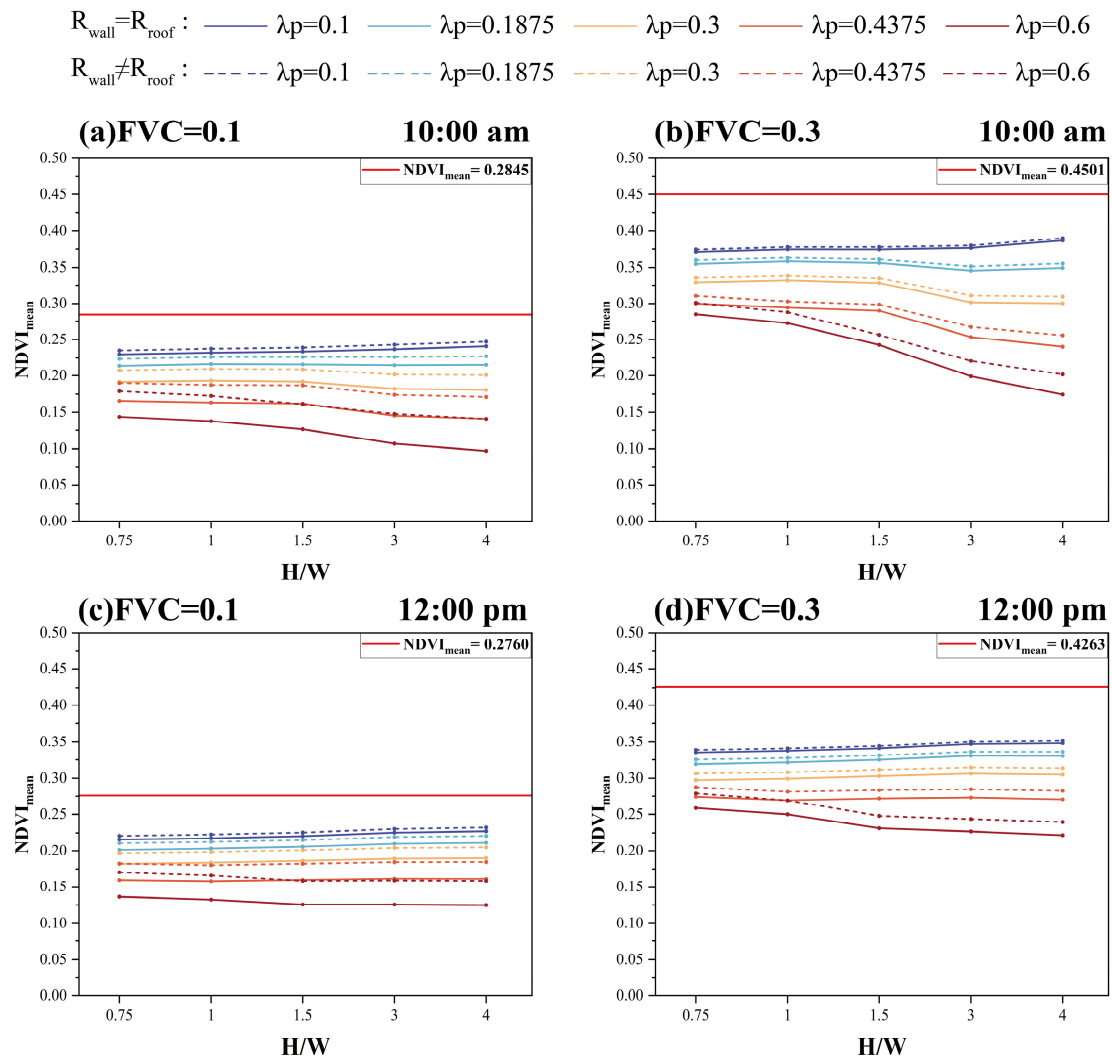


Figure 3. NDVI changes with H/W: (a) with shadow conditions and FVC = 0.1; (b) with shadow conditions and FVC = 0.3; (c) under sunlit conditions and FVC = 0.1; (d) under sunlit conditions and FVC = 0.3. The red straight lines represent the NDVI_{mean} values of the corresponding vegetation-only scenes (Table 5).

Likewise, with the same FVC, the NDVI_{mean} under shadow conditions is mostly higher than that under sunlit conditions, and it is closer to the red straight line for the scenes with shadow conditions than for the scenes under sunlit conditions (Figure 3). Generally, the NDVI with shadow conditions is higher than that under sunlit conditions. This may be because the fraction of diffuse solar irradiance reflected by adjacent pixels is a large fraction of the total irradiance onto the observed pixel under shadow conditions. The reflectance of both vegetation and buildings in the red band is lower than that in the NIR band. Thus, the scattered radiance and irradiance in the red bands is lower than in the NIR band. This makes the NDVI under shadow conditions higher than under sunlit conditions. With the increase of the H/W, the SVF decreases and reduces the NIR radiance reflected to the sky. This makes the NDVI decrease with the H/W. Thus, when the H/W, λ_p , and FVC are all high, the reduced radiance within pixels due to scattering and absorption is higher than the added radiance from adjacent pixels. Then, the NDVI_{mean} under shadow conditions is lower than that under sunlit conditions (Figure 3b,d).

3.2. Effects of Building Materials on NDVI

To investigate the effects of building materials on the NDVI, the $NDVI_{mean}$ of scenes with the same H/W , λ_p , and materials reflectance are presented in Figures 4 and 5. Table 2 illustrates the reflectance of different R_{walls} in the red and near-infrared bands. The characteristics of homogeneous and heterogeneous buildings are explained in Section 2.1.1.

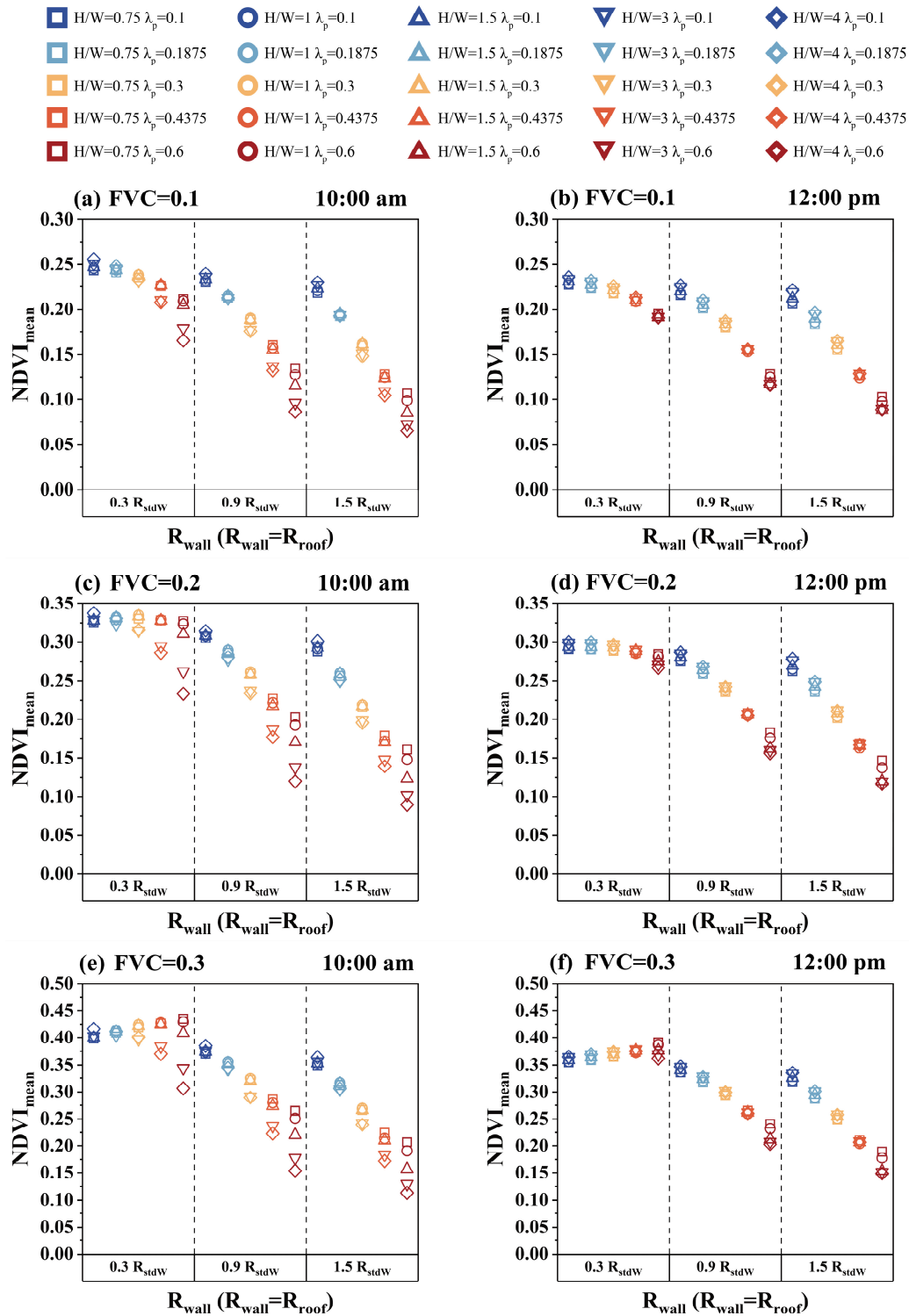


Figure 4. NDVI changes with building density and building height when building materials are the same: (a) with shadow conditions and $FVC = 0.1$; (b) under sunlit conditions and $FVC = 0.1$; (c) with shadow conditions and $FVC = 0.2$; (d) under sunlit conditions and $FVC = 0.2$; (e) with shadow conditions and $FVC = 0.3$; and (f) under sunlit conditions and $FVC = 0.3$.

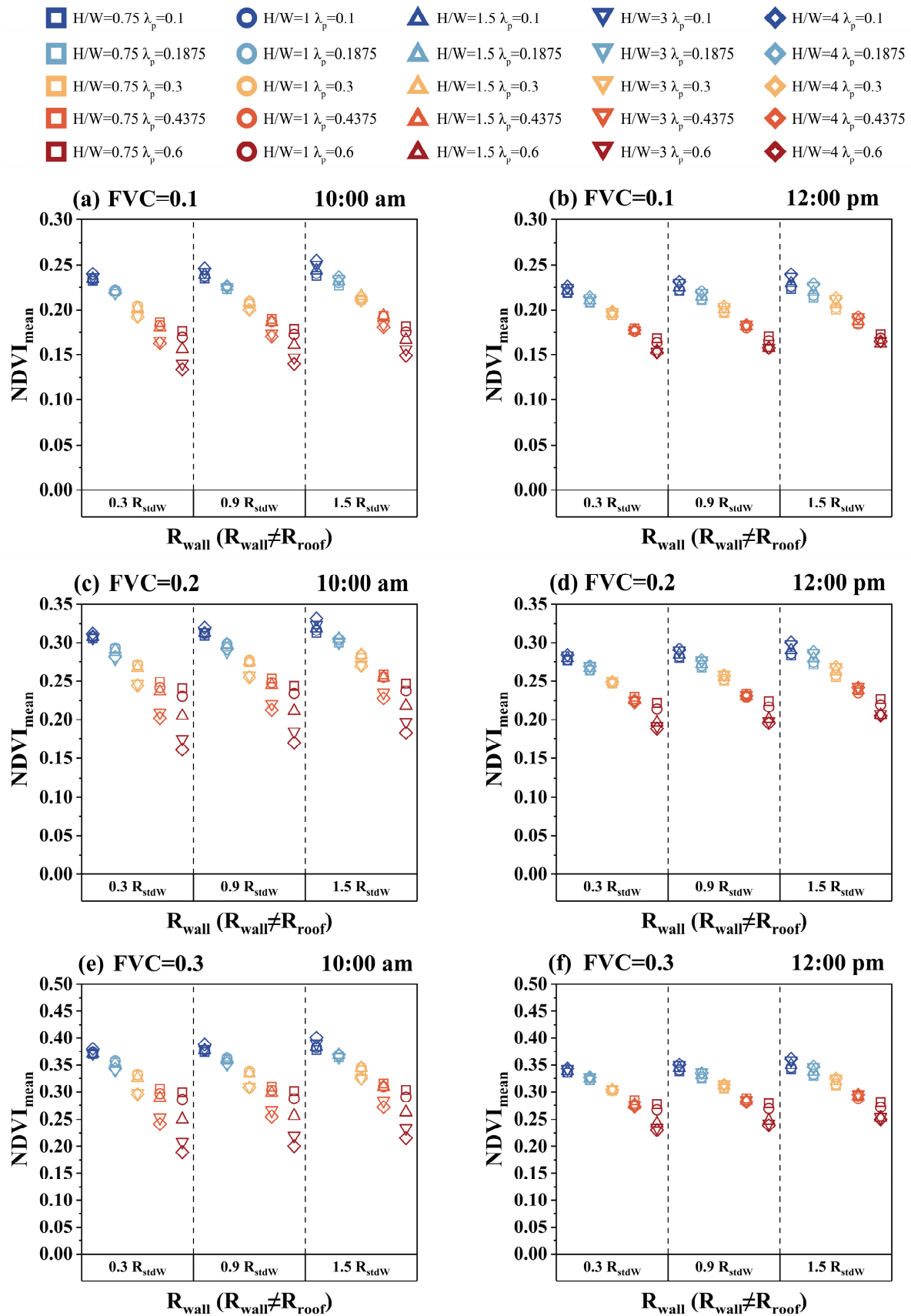


Figure 5. NDVI changes with building density and building height when building materials are heterogeneous: (a) with shadow conditions and FVC = 0.1; (b) under sunlit conditions and FVC = 0.1; (c) with shadow conditions and FVC = 0.2; (d) under sunlit conditions and FVC = 0.2; (e) with shadow conditions and FVC = 0.3; and (f) under sunlit conditions and FVC = 0.3.

Figure 4 displays the $NDVI_{mean}$ values for mixed scenes with homogeneous buildings, wherein a notable negative correlation is observed between the $NDVI_{mean}$ and the R_{wall} . Specifically, for scenes under shadow conditions (Figure 4a,c,e), when the building density is low ($\lambda_p = 0.1, 0.1875$), the NDVI is almost unaffected by the H/W and shows a highly overlapping scatter. In contrast, when the building density is high ($\lambda_p = 0.3, 0.4375, 0.6$), there is a negative correlation between the NDVI and both the H/W and R_{wall} . Moreover, when $R_{wall} = 0.3R_{stdW}$, the negative correlation between the $NDVI_{mean}$ and the building density (i.e., λ_p) gradually diminishes with the increase of the FVC, and their correlation even becomes positive when $FVC = 0.3$ (Figure 4).

Figure 5 illustrates that for scenes with heterogeneous buildings, both the $NDVI_{mean}$ and R_{wall} are weakly positively correlated. This positive correlation remains relatively consistent as the FVC increases. However, this positive correlation is marginally weaker in scenes under sunlit conditions than in scenes with shadow conditions. Only two observation times were considered for mixed scenes. This correlation is also likely to increase further if the observation time corresponds to a smaller solar elevation angle (i.e., a larger shaded area). However, at least in unshaded scenes, the effects of building materials on the NDVI are essentially negligible when the wall and roof materials are heterogeneous.

When both the scene geometry and vegetation components are the same, the $NDVI_{mean}$ of a mixed scene with heterogeneous buildings is not always greater than the $NDVI_{mean}$ of a mixed scene with homogeneous buildings. Heterogeneous buildings, however, cause the range of variation of the NDVI with the geometry to be significantly reduced (Figures 4 and 5). This suggests that heterogeneous buildings weaken the effects of geometry on the NDVI. Additionally, the difference in the $NDVI_{mean}$ between heterogeneous and homogeneous buildings becomes more pronounced as the λ_p increases. This is because a high building density increases the scattering within pixels; thus, the effects of material heterogeneity within pixels on the NDVI are higher.

3.3. Accuracy of Three FVC Estimation Methods

Figure 6 illustrates the FVC difference (dif_{FVC}) between the predefined FVC (FVC_{true}) and the FVC retrieved with the three algorithms (Equations (3)–(5)) for four types of scenes:

- Homogeneous building under shadow conditions ($R_{wall} = R_{roof}$, 10:00 am)
- Homogeneous building under sunlit conditions ($R_{wall} = R_{roof}$, 12:00 pm)
- Heterogeneous building under shadow conditions ($R_{wall} \neq R_{roof}$, 10:00 am)
- Heterogeneous building under sunlit conditions ($R_{wall} \neq R_{roof}$, 12:00 pm)

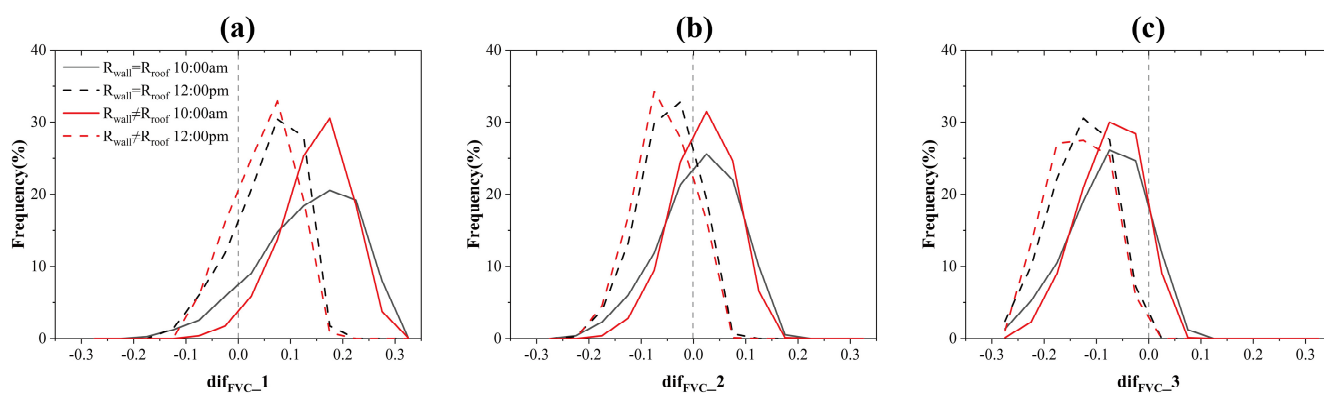


Figure 6. Histograms of the FVC difference (dif_{FVC}) between the predefined FVC_{true} and the FVC estimates by three algorithms: (a) FVC estimates based on Equation (3); (b) FVC estimates based on Equation (4); and (c) FVC estimates based on Equation (5).

The findings indicate that Equation (3) gave the least accurate estimates of the FVC, with the FVC typically exceeding the FVC_{true} . The impact of the building material heterogeneity on the FVC estimation accuracy is inconsequential under sunlit conditions. Under shadow conditions, the dif_{FVC_1} for heterogeneous buildings is predominantly concentrated between 0.1 and 0.2, whereas the dif_{FVC_1} for homogeneous buildings exhibits greater dispersion. Likewise, the FVC estimation accuracy for scenes under shadow conditions is markedly inferior to that for scenes under sunlit conditions, indicating that shadows result in a considerable overestimation of the FVC (Figure 6a).

The FVC estimation accuracy based on Equation (4) is the highest, with minimal differences among the four types of scenes, indicating that the method is more robust and less susceptible to fluctuations in accuracy due to the diverse urban scenes (Figure 6b).

The FVC estimation accuracy based on Equation (5) is moderate, with the smallest differences among the four types of scenes and a general underestimation of the FVC. However, the FVC estimation accuracy for scenes under shadow conditions is slightly higher than that for scenes under sunlit conditions, and there is minimal difference in the FVC estimation accuracy for homogeneous and heterogeneous buildings (Figure 6c).

Table 6 demonstrates that FVC_2 performs best ($0.060 \leq RMSE \leq 0.077$, $-0.057 \leq BIAS \leq 0.018$), with the exception of scenes with heterogeneous buildings under sunlit conditions, where FVC_1 is the best algorithm. There is a minimal difference in accuracy between FVC_1 ($0.074 \leq RMSE \leq 0.163$, $0.048 \leq BIAS \leq 0.148$) and FVC_3 ($0.095 \leq RMSE \leq 0.146$, $-0.135 \leq BIAS \leq -0.074$), but the difference in accuracy among the four types of scenes for FVC_1 is significantly greater.

Table 6. Accuracy of different FVC estimation algorithms.

| | FVC_1 | FVC_2 | FVC_3 | |
|------|--|--|--|--|
| RMSE | 0.128 | 0.070 | 0.124 | |
| BIAS | 0.096 | −0.003 | −0.104 | |
| | FVC_1 | | | |
| | R _{wall} = R _{roof} (10:00 am) | R _{wall} = R _{roof} (12:00 pm) | R _{wall} ≠ R _{roof} (10:00 am) | R _{wall} ≠ R _{roof} (12:00 pm) |
| RMSE | 0.163 | 0.084 | 0.162 | 0.074 |
| BIAS | 0.133 | 0.057 | 0.148 | 0.048 |
| | FVC_2 | | | |
| | R _{wall} = R _{roof} (10:00 am) | R _{wall} = R _{roof} (12:00 pm) | R _{wall} ≠ R _{roof} (10:00 am) | R _{wall} ≠ R _{roof} (12:00 pm) |
| RMSE | 0.074 | 0.073 | 0.060 | 0.077 |
| BIAS | 0.010 | −0.049 | 0.018 | −0.057 |
| | FVC_3 | | | |
| | R _{wall} = R _{roof} (10:00 am) | R _{wall} = R _{roof} (12:00 pm) | R _{wall} ≠ R _{roof} (10:00 am) | R _{wall} ≠ R _{roof} (12:00 pm) |
| RMSE | 0.107 | 0.141 | 0.095 | 0.146 |
| BIAS | −0.079 | −0.129 | −0.074 | −0.135 |

Figure 7 illustrates the absolute value of the difference between the predefined FVC_{true} and the FVC estimates, i.e., the (dif_{FVC_1} , dif_{FVC_2} , dif_{FVC_3} , and $|dif_{FVC_1}|$, $|dif_{FVC_2}|$, $|dif_{FVC_3}|$) defined in Section 2.2. A smaller $|dif_{FVC}|$ indicates a higher FVC estimation accuracy. The building density of the scenes, from left to right, increases in a sequential manner across all the intervals with equal H/W. The specific parameters of scenes “S1~S25” are shown in Figure A1.

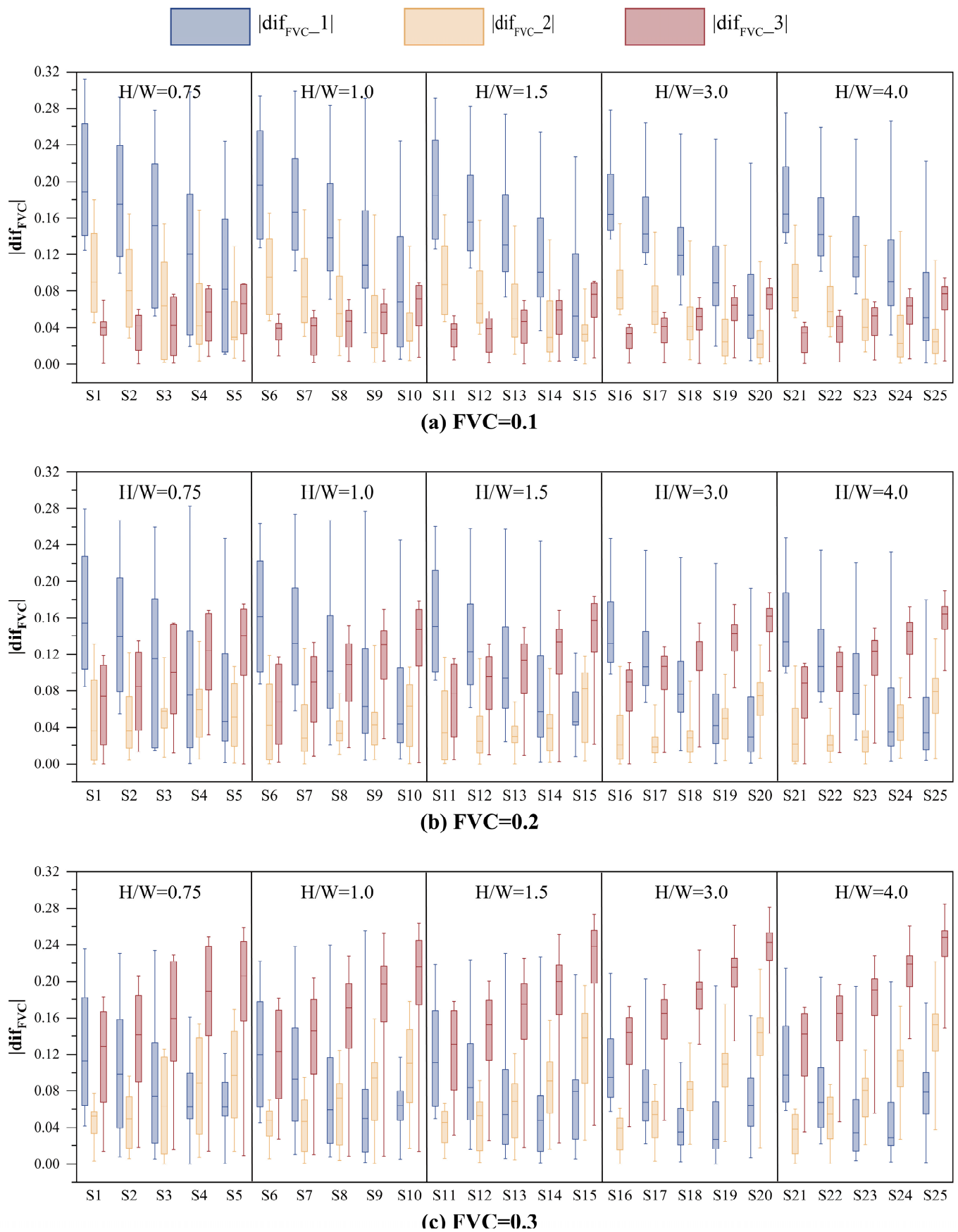


Figure 7. Range of absolute values ($|dif_{FVC}|$) of the FVC difference between the predefined FVC_{true} and the FVC estimates by three types of algorithms in urban scenes S1 to S25 (Figure A1): (a) $FVC = 0.1$; (b) $FVC = 0.2$; and (c) $FVC = 0.3$.

It can be observed that when $FVC = 0.1$, the estimation accuracies by FVC_1 and FVC_2 are significantly and positively correlated with the building density. In contrast, the estimation accuracies by FVC_3 display a negative correlation with the building density (Figure 7a). When $FVC = 0.2$, the estimation accuracy by FVC_1 remains positively correlated with the building density, and the estimation accuracy by FVC_3 continues to exhibit a negative correlation. However, the relationship between the estimation accuracy of FVC_2 and the building density is not yet clear when $H/W = 0.75$. As H/W increases, the FVC estimation accuracy initially exhibits a positive correlation with the building density, followed by a negative correlation, with the turning point in this correlation shifting toward a lower building density. This suggests that the estimation accuracy of FVC_2 for medium vegetation cover is significantly reduced when both the building height and density are high (Figure 7b). When $FVC = 0.3$, the estimation accuracy of both FVC_2 and FVC_3 is negatively correlated with the building density, and the relationship between the estimation accuracy of FVC_1 and the building density is analogous to the pattern observed in FVC_2 when $FVC = 0.2$ (Figure 7c).

The longer boxes indicate that the FVC estimation accuracy is significantly influenced by the sunlight conditions and building materials, suggesting that the algorithm is unstable (Figure 7). As the FVC increases, the stability of FVC_1 is shown to be “high FVC > medium FVC > low FVC”. In contrast, the stability of FVC_2 is relatively constant, while the stability of FVC_3 is shown to be “low FVC > medium FVC > high FVC”. Additionally, as for the FVC estimation accuracy (Figure 7), FVC_1 exhibited a trend of “high FVC > medium FVC > low FVC,” FVC_2 demonstrated a pattern of “medium FVC > low FVC > high FVC”, and FVC_3 displayed a pattern of “low FVC > medium FVC > high FVC”.

It is therefore proposed that the estimation accuracy of FVC_3 is optimal and stable when the vegetation cover is low. In the case of medium vegetation cover, the estimation accuracy of FVC_2 is optimal and stable. However, in situations where the vegetation cover is high, it is recommended that FVC_2 is used for a low building density ($\lambda_p < 0.3$) and FVC_1 for a high building density ($\lambda_p \geq 0.3$).

3.4. Validation

Table 7 and Figure 8 illustrate three FVC algorithms, which collectively demonstrate a significant improvement in the FVC estimation accuracy through the application of Equation (9) for correction. The accuracy rankings of the three FVC algorithms in our study area are as follows: FVC_3 , FVC_2 , and FVC_1 (Table 7). To validate the results in Section 3.3, the “true FVC” values (refer to Section 2.4) were classified into four intervals: 0~0.1, 0.1~0.2, 0.2~0.3, and greater than 0.3, respectively (Table 8). Regardless of the algorithm employed to calculate the “true FVC”, a greater than 50% pixel count within the “true FVC” range of 0~0.1 indicates low overall vegetation cover in the study area. This is the same as the simulation results in Section 3.3, “The accuracy of the three algorithms performs as $FVC_3 > FVC_2 > FVC_1$ when $0 < FVC \leq 0.1$ ” (Figure 7a).

Table 7. Accuracy of different FVC estimation algorithms before and after correction using Equation (9).

| | | FVC_1 | FVC_2 | FVC_3 |
|-------------------|------|---------|---------|---------|
| Before Correction | RMSE | 0.249 | 0.183 | 0.169 |
| | BIAS | 0.205 | 0.144 | 0.098 |
| After Correction | RMSE | 0.164 | 0.114 | 0.087 |
| | BIAS | 0.119 | 0.080 | 0.040 |

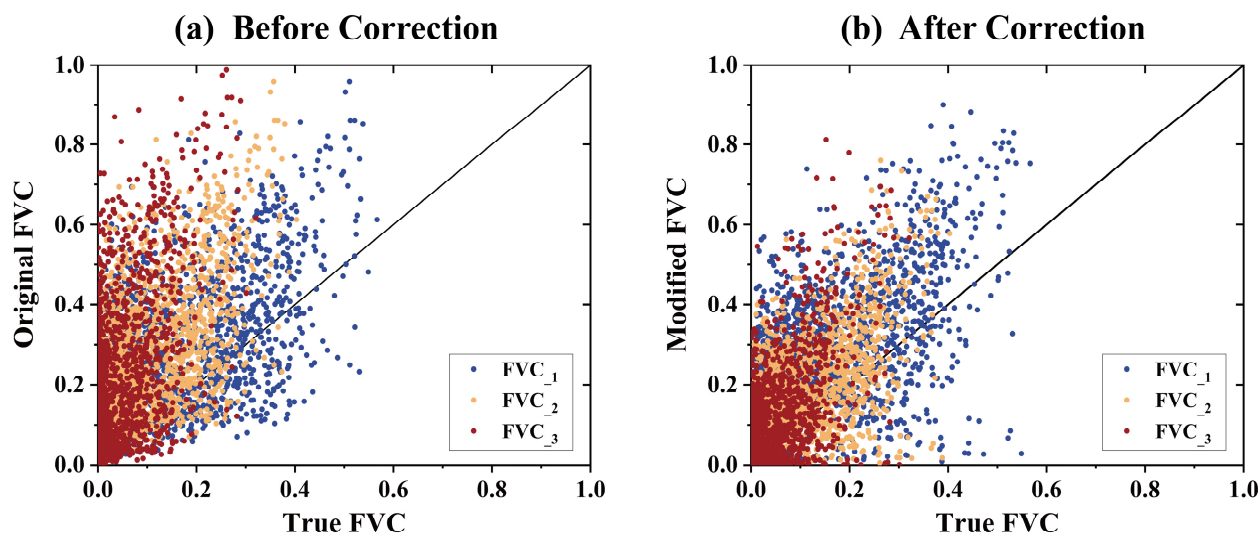


Figure 8. Accuracy of different FVC estimation algorithms before and after correction using Equation (9), where the axis labels are predefined in Section 2.4: (a) before correction; and (b) after correction.

Table 8. Accuracy of different FVC estimation algorithms before and after correction using Equation (9) at different “true FVC” intervals.

| | | 0 < True FVC ≤ 0.1 | | |
|-------------------|------|----------------------|------------------|----------------------|
| | | FVC ₁ | FVC ₂ | FVC ₃ |
| Before Correction | RMSE | 0.250 | 0.177 | 0.142 |
| | BIAS | 0.216 | 0.151 | 0.082 |
| After Correction | RMSE | 0.173 | 0.117 | 0.080 |
| | BIAS | 0.139 | 0.102 | 0.030 |
| | | 0.1 < True FVC ≤ 0.2 | | 0.1 < True FVC ≤ 0.3 |
| | | FVC ₁ | FVC ₂ | FVC ₃ |
| Before Correction | RMSE | 0.123 | 0.073 | 0.257 |
| | BIAS | 0.103 | 0.050 | 0.229 |
| After Correction | RMSE | 0.089 | 0.050 | 0.173 |
| | BIAS | 0.053 | 0.023 | 0.130 |
| | | 0.2 < True FVC ≤ 0.3 | | |
| | | FVC ₁ | FVC ₂ | |
| Before Correction | RMSE | 0.104 | 0.200 | |
| | BIAS | 0.068 | 0.159 | |
| After Correction | RMSE | 0.084 | 0.124 | |
| | BIAS | 0.035 | 0.105 | |

Given the limited number of pixels with “True FVC” values exceeding 0.3 within the study area, we only focused on the performance of each FVC estimation algorithm when $0 < \text{“True FVC”} \leq 0.3$. The majority of the “True FVC” values of FVC₃ are situated within the range of 0.1~0.2, and we combined the pixels in the intervals of 0.1~0.2 and 0.2~0.3. A comparative analysis of these data (Table 8) reveals that the performance of the three FVC algorithms (Figure 7) and selection recommendations (Section 3.3) is almost the same as our validation results (Table 8). The application of Equation (9) for correction, followed by the selection of an appropriate FVC algorithm for the study area based on the aforementioned recommendations, may be an effective method for retrieving the urban vegetation abundance, as evidenced by the minimum value of the RMSE being 0.050 (Table 8).

4. Discussion

The NDVI is one of the most widely used vegetation indices for assessing the biophysical and biochemical characteristics of vegetation, such as the fractional vegetation cover

(FVC), leaf area index (LAI), chlorophyll content, gross primary productivity (GPP), and fraction of absorbed photosynthetically active radiation (FAPAR) [46]. It is affected by the background reflectance, solar radiation and observation geometry, since these factors alter the reflectance, and by the bidirectional reflectance distribution function (BRDF) [46]. The urban geometry and materials are highly heterogeneous and may also affect the reflectance in different bands. To address this issue, it is important to consider the effects of the urban geometry and background reflectance on the NDVI. In our study, we examined the impact of the urban geometry and background materials on the NDVI and its application in estimating the vegetation fractional abundance in urban areas. Experiments were carried out by applying the DART model to a large set of synthetic urban scenes to calculate the NDVI and FVC separately.

We attempted to interpret the results in Section 3.1 by sampling pixels (Figure 9) and comparing the NDVI values of the components (Figure 10) within mixed scenes. Figure 9c shows a schematic of the sampled pixels, which were randomly selected from a mixed scene. The reflectance in the red and near-infrared bands was taken at 10 am (i.e., shaded) and 12 pm (i.e., unshaded) for these sampling pixels. The findings indicate that shadows have a considerable impact on the NDVI, as evidenced by the presence of elevated NDVI values in the shaded regions. Sunlit areas mainly receive direct solar radiation, while shaded areas receive diffuse solar radiation from the sky and scattered radiation from adjacent pixels. Due to the absence of direct solar radiation in shaded areas, the reflectance in the red and near-infrared bands decreases significantly, the reflectance in the red band decreases faster (Figure 9a), and the ratio of $(\rho_{\text{NIR}} - \rho_{\text{Red}})/(\rho_{\text{NIR}} + \rho_{\text{Red}})$ increases (Figure 9b). The variable k in Figure 9b represents the slope of the line connecting each data point with the origin and the NDVI value of the pixel. A straightforward calculation reveals that for pixels observed from 12 pm to 10 am, the average $\rho_{\text{NIR}} + \rho_{\text{Red}}$ value decreases from 0.597 to 0.139 and the average $\rho_{\text{NIR}} - \rho_{\text{Red}}$ value decreases from 0.206 to 0.094 (Figure 9b). The sum value of the two bands decreases at a faster rate than the difference value. The NDVI values for pixels at 12 pm range from 0.216 to 0.906, while for pixels at 10 am, they range from 0.288 to 0.921. For shadow areas, the irradiance from adjacent pixels cannot be neglected, and the NIR band is scattered heavier than the red band, as evidenced by the reflectance in the near-infrared band being always greater than that in the red band (Figure 9a); thus, the estimated NDVI in shadow areas under these illumination conditions may be higher. This means that the NDVI of the pixels in shadow areas is overestimated to some extent.

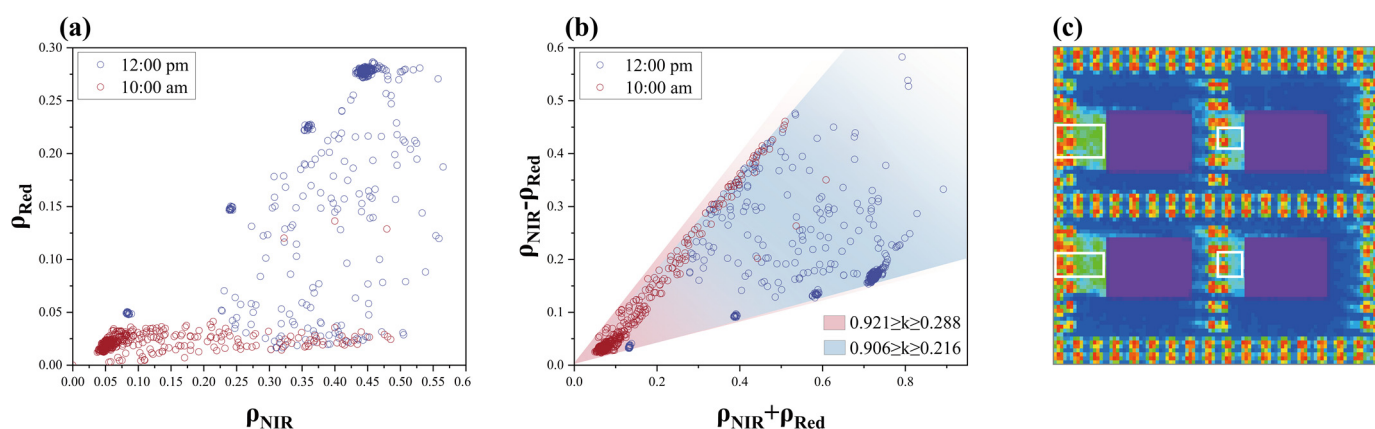


Figure 9. Ground component pixels of the scene under sunlit and shadow conditions: (a) reflectance in the red and near-infrared bands; (b) sum and difference of reflectance in the red and near-infrared bands; and (c) schematic diagram of sampled pixels (inside the white boxes).

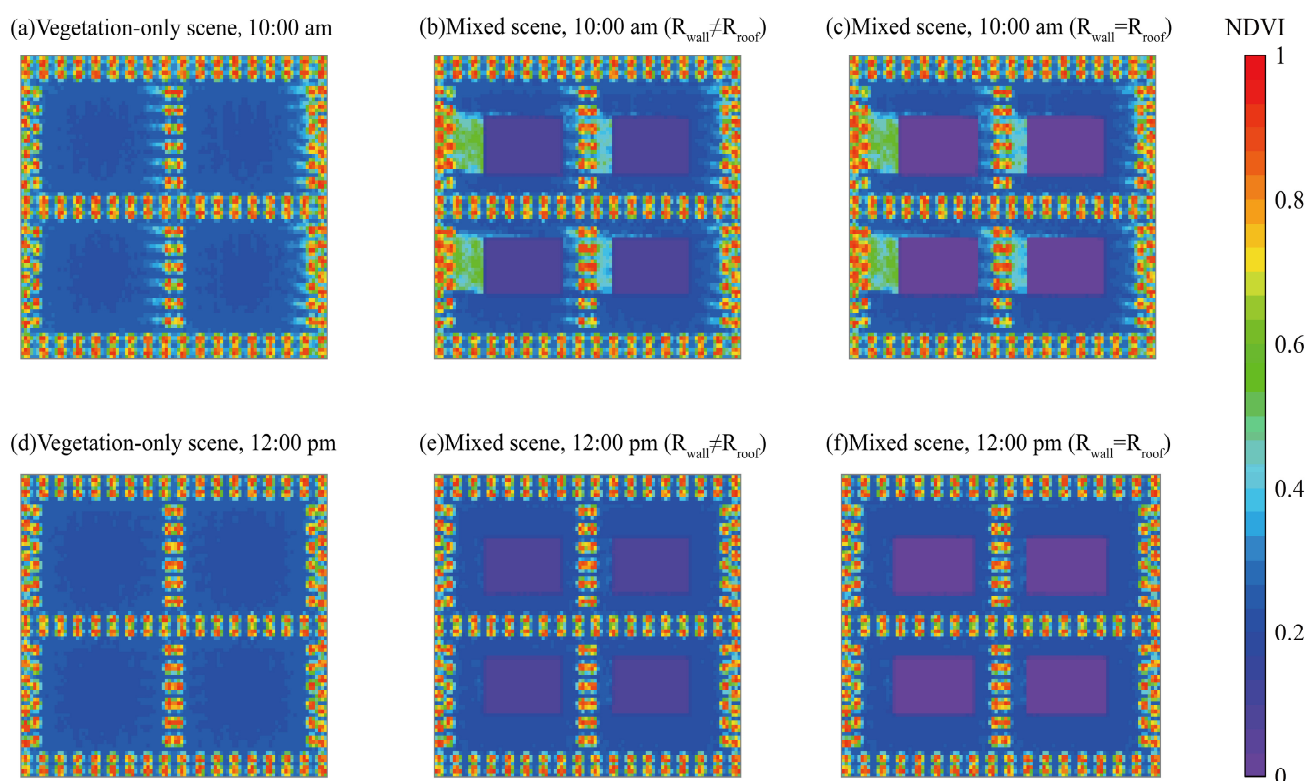


Figure 10. NDVI derived in the DART model for four types of scenes: (a) vegetation-only scene with shadow conditions; (b) mixed scene with heterogeneous buildings and shadow conditions; (c) mixed scene with homogenous buildings and shadow conditions; (d) vegetation-only scene under sunlit conditions; (e) mixed scene with heterogeneous buildings under sunlit conditions; and (f) mixed scene with homogenous buildings under sunlit conditions.

Figure 10 provides a more illustrative representation of this phenomenon. In scenes with shadow conditions, the NDVI is relatively high in the shaded region caused by buildings (Figure 10a–c). The NDVI of the shaded region at the periphery of the building complex is found to be higher than that of the shaded region within the building complex, and the NDVI of vegetation within the shaded region is slightly higher than in the sunlit vegetation pixels (Figure 10b,c). Irrespective of the time of the day, the NDVI of the building pixels is approximately zero, whereas the NDVIs of the corresponding pixels in vegetation-only scenes are markedly higher than that of the building pixels in mixed scenes (Figure 10). The entire scene is treated as a single pixel, with the signal received by a spectro-radiometer representing the sum of the component signals within the target pixel. This allows us to define the mean NDVI of the urban mixed scene as the NDVI of the pixel. This explains the experimental result (Section 3.1) that for a certain H/W, an elevated building density (i.e., an expanded area where NDVI is proximate to zero) is associated with a diminished NDVI, and shadow may reduce the bias of NDVI estimates in urban areas. Our conclusion was also validated with satellite data provided by Bhang and Lee et al. [47]. A similar conclusion was obtained with linear spectral mixing models by Jiang et al. [48], despite the fact that this conclusion was derived from a mixed pixel comprising three components, i.e., soil, vegetation, and the projection of vegetation on soil. In our scenes, the presence of buildings is also considered. This further suggests that the presence of shadows increases the NDVI, both for the mixed pixels consisting of simple soil and vegetation and for mixed pixels in urban areas.

As illustrated in Figure 10, the NDVIs of the urban mixed pixels are predominantly influenced by the shaded area and the building area within the pixels. In Section 3.1, at a

low building density ($\lambda_p = 0.1, 0.1875$), the NDVI values of the mixed scenes observed at 10 am demonstrate an increase with the H/W, whereas those observed at 12 pm exhibit no correlation with the H/W. This suggests that shadows exert a predominant influence on the NDVI when buildings are open. At high building densities ($\lambda_p = 0.3, 0.4375, 0.6$), regardless of the time of day when mixed scenes are observed, the NDVI decreases despite the fact that the H/W is increasing at this time, which is likely to be due to the reduced NIR reflectance when the SVF decreases as the H/W increases. Urban FVC estimation accuracy is inevitably affected by the misestimation of the NDVI in urban mixed pixels due to shadows and building areas. The systematic study of biomass estimation by means of vegetation indices conducted by Peddle et al. [49] revealed that the information that vegetation indices can provide is at the pixel scale, and the use of sub-pixel information (e.g., the area of shaded regions in a pixel) was found to greatly improve the accuracy of biomass estimates. Consequently, there may be a promising approach to enhance the reliability of the NDVI in satellite imagery for applications in urban areas, which entails the construction of a model that can determine the area of shaded regions and buildings by utilizing higher spatial resolution building density and height data for a given solar altitude angle and observation angle. This would then allow for the implementation of different levels of adjustment to the NDVI in urban areas for high-accuracy inversion of the FVC.

Interestingly, our findings indicate that the NDVI is affected by the reflectance of building materials when shadow components are present within a mixed scene comprising heterogeneous buildings (Section 3.2). Adeline et al. [50] provided compelling evidence supporting our results by presenting the data of numerical experiments using the radiative transfer model Amartis v2 [51]. Their scenario was designed to reflect a heterogeneous and shaded building with walls and a roof, a configuration similar to some of our own scenario setups. In our scenes, we gave additional consideration to the absence of shadows and found that the heterogeneous building had essentially negligible effects on the NDVI under sunlit condition. However, further verification of this observation with data derived by other models, field measurements, and satellite data is required.

The relationships between the urban geometry and background reflectance and the NDVI and FVC are proposed in this study. The relationships mentioned in our study are derived from a simple symmetric urban scene, assuming that the 3D urban geometry and background reflectance are the main factors controlling the NDVI. Indeed, the urban geometry within mixed pixels in actual urban scenes is often more complex. This study considered only one atmospheric condition. However, atmospheric conditions are variable, and the signals received by the sensors are often mixed with atmospheric effects and signals from adjacent pixels, which remain to be investigated. Likewise, the magnitudes of the building heights and densities selected in this paper are based on the general characteristics of typical urban scenes, with a maximum building height of 40 m. In some special urban clusters with ultra-high building heights, such as Hong Kong, the building heights generally exceed 100 m, with very large H/W values, which may necessitate further research.

Despite the limitations of our experiments, our study has yielded valuable conclusions. By quantifying the geometry and background reflectance of a typical urban area by applying the DART model, we have explored the general pattern of the NDVI in urban areas and provided some useful suggestions for addressing the scientific challenge of “how to estimate FVC using NDVI in urban areas” (Section 3.3). This will assist future research in achieving a deeper understanding of the impact of the urban 3D environment on red and near-infrared bands’ reflectance, as well as in the selection of urban FVC inversion methods when using space-borne multi-spectral data.

Likewise, we found that the results obtained from the simulation data based on the DART model overlapped with many previous studies and were also the same as the satellite data validation results (Section 3.4), which further demonstrated the effectiveness of the model in simulating urban areas. We also discovered that many conclusions that are challenging to ascertain from satellite data can be derived from simulated data. The radiance of the mixed pixels received by a remote-imaging radiometer is influenced by a multitude of factors, which can distort the relationships between them. However, this issue can be mitigated to some extent by the use of simulated data, i.e., better controlled experiments. It can therefore be argued that simulated data are a valuable complement to satellite data in exploring relationships that are difficult to establish using satellite imagery only. However, it should be noted that the results obtained through numerical experiments still require validation through the use of other models, in situ measurements and satellite data to obtain more rigorous scientific answers.

5. Conclusions

This study simulated various urban scenes using the DART model to investigate the impact of the urban geometry and background reflectance on the NDVI, and it also explored how to select a suitable FVC estimation method using the NDVI to estimate the urban vegetation component abundance when using space-borne multi-spectral data. Some insightful findings are the following:

- (1) When building wall and roof materials are heterogeneous, the effects of material differences on the NDVI are essentially negligible if there is minimal or no shadow components within the mixed urban pixels in space-borne satellite images. Conversely, when shadow components are present, the effects are non-negligible (Section 3.2).
- (2) The NDVI of an urban mixed pixel is determined by a combination of the pixel's internal shadow components, which increase the NDVI, and the proportion of building area, which decreases the NDVI (Section 4). To enhance the accuracy of NDVI-based FVC algorithms, the ideal NDVI exclusively reflects the vegetation component and is not affected by buildings. When the H/W increases, the NDVI will be underestimated because of the reduced NIR reflectance caused by multiple scattering and reabsorption (Section 3.1), which will affect the accuracy of NDVI-based FVC algorithms.
- (3) Introducing the SVF to correct the NDVI has proven an effective means of estimating the urban FVC (Section 3.4). Further improvements in the FVC estimation accuracy may be achieved by combining this approach with the characteristics of the study area and the FVC algorithm selection recommendations (Section 3.3).

Author Contributions: Conceptualization, J.Y., Y.X. and H.C.H.; methodology, J.Y.; software, W.X.; validation, W.X.; investigation, W.X.; data curation, W.X. and L.F.; writing—original draft preparation, W.X., L.F. and J.Y.; writing—review and editing, J.Y., Y.X., H.C.H., M.M., R.L. and M.S.W.; visualization, W.X.; supervision, J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Grants by the National Natural Science Foundation of China (42271345, 42071394) and National College Students Innovation and Entrepreneurship Training Program (S202411078029, 202311078005). Massimo Menenti is thankful for the support from the MOST High Level Foreign Expert program (Grant No. G2022055010L) and the Chinese Academy of Sciences President's International Fellowship Initiative (Grant No. 2020VTA0001). M.S. Wong thanks the funding support from the General Research Fund (Grant No. 15603920 and 15609421), and the Collaborative Research Fund (Grant No. C5062-21GF) from the Research Grants Council, Hong Kong, China.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

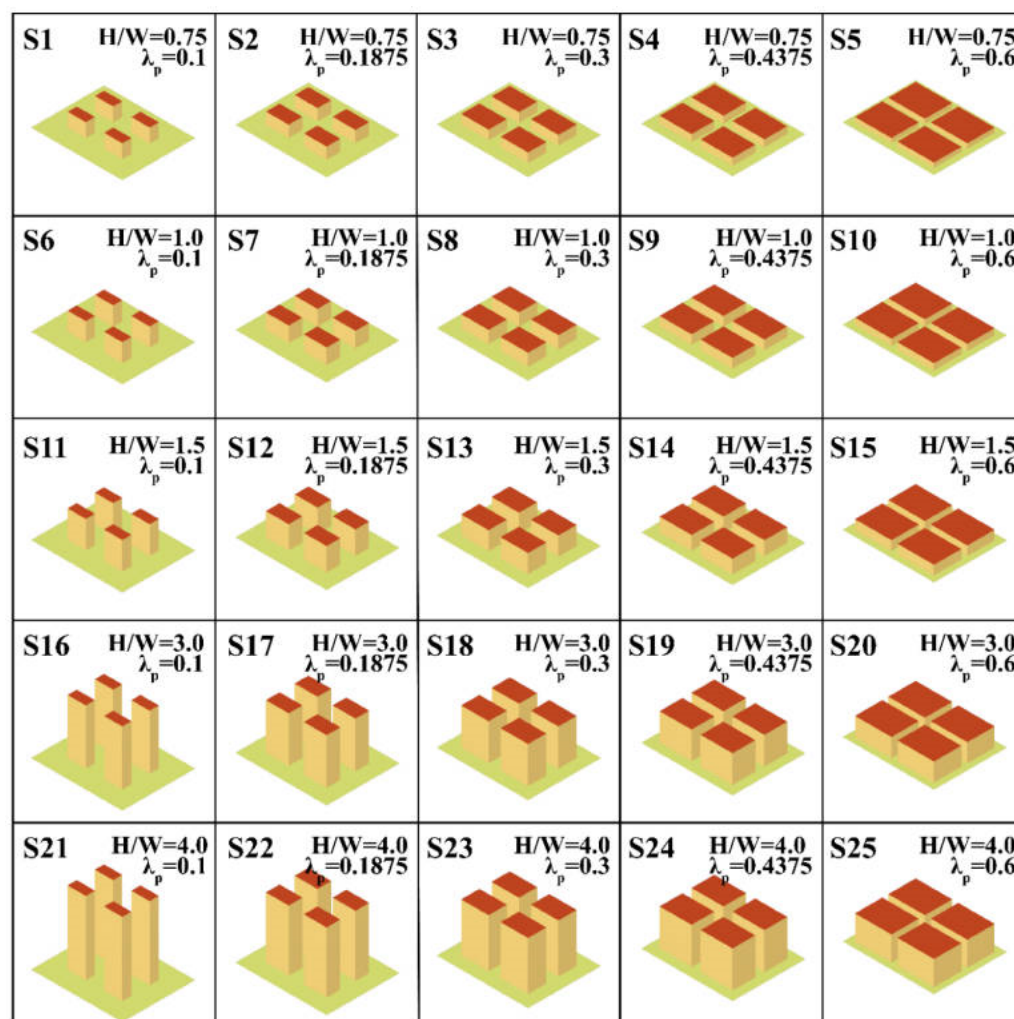


Figure A1. Illustrates the display and numbering of different building scenes.

References

- Wu, S.; Yu, W.; An, J.; Lin, C.; Chen, B. Remote sensing of urban greenspace exposure and equality: Scaling effects from greenspace and population mapping. *Urban For. Urban Green.* **2023**, *90*, 128136. [\[CrossRef\]](#)
- Chen, W.; Huang, H.; Dong, J.; Zhang, Y.; Tian, Y.; Yang, Z. Social functional mapping of urban green space using remote sensing and social sensing data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 436–452. [\[CrossRef\]](#)
- Shi, Q.; Liu, M.; Marinoni, A.; Liu, X. UGS-1m: Fine-grained urban green space mapping of 31 major cities in China based on the deep learning framework. *Earth Syst. Sci. Data* **2023**, *15*, 555–577. [\[CrossRef\]](#)
- Yang, J.; He, Y.; Oguchi, T. An endmember optimization approach for linear spectral unmixing of fine-scale urban imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *27*, 137–146. [\[CrossRef\]](#)
- Zhang, Y.; Harris, A.; Balzter, H. Characterizing fractional vegetation cover and land surface temperature based on sub-pixel fractional impervious surfaces from Landsat TM/ETM+. *Int. J. Remote Sens.* **2015**, *36*, 4213–4232. [\[CrossRef\]](#)
- Nguyen, T. Fractional Vegetation Cover Change Detection in Megacities Using Landsat Time-Series Images: A Case Study of Hanoi City (Vietnam) During 1986–2019. *Geogr. Environ. Sustain.* **2019**, *12*, 175–187. [\[CrossRef\]](#)
- Gao, L.; Wang, X.; Johnson, B.A.; Tian, Q.; Wang, Y.; Verrelst, J.; Mu, X.; Gu, X. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 364–377. [\[CrossRef\]](#)
- Cheng, Y.; Wang, W.; Ren, Z.; Zhao, Y.; Liao, Y.; Ge, Y.; Wang, J.; He, J.; Gu, Y.; Wang, Y.; et al. Multi-scale Feature Fusion and Transformer Network for urban green space segmentation from high-resolution remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103514. [\[CrossRef\]](#)
- Li, T.; Meng, Q. A mixture emissivity analysis method for urban land surface temperature retrieval from Landsat 8 data. *Landsc. Urban Plan.* **2018**, *179*, 63–71. [\[CrossRef\]](#)

10. Li, Z.-L.; Wu, H.; Wang, N.; Qiu, S.; Sobrino, J.A.; Wan, Z.; Tang, B.-H.; Yan, G. Land surface emissivity retrieval from satellite data. *Int. J. Remote Sens.* **2012**, *34*, 3084–3127. [[CrossRef](#)]
11. Sobrino, J.A.; Jiménez-Muñoz, J.C.; Sòria, G.; Romaguera, M.; Guanter, L.; Moreno, J.; Plaza, A.; Martínez, P. Land surface emissivity retrieval from different VNIR and TIR sensors. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 316–327. [[CrossRef](#)]
12. Weng, Q.; Lu, D.; Schubring, J. Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote Sens. Environ.* **2004**, *89*, 467–483. [[CrossRef](#)]
13. Bannari, A.; Morin, D.; Bonn, F.; Huete, A.R. A review of vegetation indices. *Remote Sens. Rev.* **1995**, *13*, 95–120. [[CrossRef](#)]
14. Zeng, Y.; Hao, D.; Huete, A.; Dechant, B.; Berry, J.; Chen, J.M.; Joiner, J.; Frankenberg, C.; Bond-Lamberty, B.; Ryu, Y.; et al. Optical vegetation indices for monitoring terrestrial ecosystems globally. *Nat. Rev. Earth Environ.* **2022**, *3*, 477–493. [[CrossRef](#)]
15. Kotthaus, S.; Smith, T.E.L.; Wooster, M.J.; Grimmond, C.S.B. Derivation of an urban materials spectral library through emittance and reflectance spectroscopy. *ISPRS J. Photogramm. Remote Sens.* **2014**, *94*, 194–212. [[CrossRef](#)]
16. Yang, S.; Li, S.; Zhang, B.; Yu, R.; Li, C.; Hu, J.; Liu, S.; Cheng, E.; Lou, Z.; Peng, D. Accurate estimation of fractional vegetation cover for winter wheat by integrated unmanned aerial systems and satellite images. *Front. Plant Sci.* **2023**, *14*, 1220137. [[CrossRef](#)]
17. Zhao, J.; Li, J.; Liu, Q.; Zhang, Z.; Dong, Y. Comparative Study of Fractional Vegetation Cover Estimation Methods Based on Fine Spatial Resolution Images for Three Vegetation Types. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
18. Mu, X.; Zhao, T.; Ruan, G.; Song, J.; Wang, J.; Yan, G.; McVicar, T.R.; Yan, K.; Gao, Z.; Liu, Y.; et al. High Spatial Resolution and High Temporal Frequency (30-m/15-day) Fractional Vegetation Cover Estimation over China Using Multiple Remote Sensing Datasets: Method Development and Validation. *J. Meteorol. Res.* **2021**, *35*, 128–147. [[CrossRef](#)]
19. Liu, Q.; Zhang, T.; Li, Y.; Li, Y.; Bu, C.; Zhang, Q.J.C.G.S. Comparative analysis of fractional vegetation cover estimation based on multi-sensor data in a semi-arid sandy area. *Chin. Geogr. Sci.* **2019**, *29*, 166–180. [[CrossRef](#)]
20. Fernández-Guisuraga, J.M.; Verrelst, J.; Calvo, L.; Suárez-Seoane, S. Hybrid inversion of radiative transfer models based on high spatial resolution satellite reflectance data improves fractional vegetation cover retrieval in heterogeneous ecological systems after fire. *Remote Sens. Environ.* **2021**, *255*, 112304. [[CrossRef](#)]
21. Lee, G.; Kim, G.; Min, G.; Kim, M.; Jung, S.; Hwang, J.; Cho, S.J.A.S. Vegetation classification in urban areas by combining UAV-based NDVI and thermal infrared image. *Appl. Sci.* **2022**, *13*, 515. [[CrossRef](#)]
22. Song, W.; Zhao, T.; Mu, X.; Zhong, B.; Zhao, J.; Yan, G.; Wang, L.; Niu, Z. Using a Vegetation Index-Based Mixture Model to Estimate Fractional Vegetation Cover Products by Jointly Using Multiple Satellite Data: Method and Feasibility Analysis. *Forests* **2022**, *13*, 691. [[CrossRef](#)]
23. Jia, K.; Liang, S.; Gu, X.; Baret, F.; Wei, X.; Wang, X.; Yao, Y.; Yang, L.; Li, Y. Fractional vegetation cover estimation algorithm for Chinese GF-1 wide field view data. *Remote Sens. Environ.* **2016**, *177*, 184–191. [[CrossRef](#)]
24. Cai, Y.; Zhang, M.; Lin, H. Estimating the urban fractional vegetation cover using an object-based mixture analysis method and Sentinel-2 MSI imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 341–350. [[CrossRef](#)]
25. Hehl, M.; Hammerle, A.; Tappeiner, U.; Leitinger, G. Determinants of urban–rural land surface temperature differences—A landscape scale perspective. *Landsc. Urban Plan.* **2015**, *134*, 33–42. [[CrossRef](#)]
26. Kim, H.W.; Kim, J.-H.; Li, W.; Yang, P.; Cao, Y. Exploring the impact of green space health on runoff reduction using NDVI. *Urban For. Urban Green.* **2017**, *28*, 81–87. [[CrossRef](#)]
27. Gastellu-Etchegorry, J.P.; Lauret, N.; Yin, T.; Landier, L.; Kallel, A.; Malenovsky, Z.; Bitar, A.A.; Aval, J.; Benhmida, S.; Qi, J.; et al. DART: Recent Advances in Remote Sensing Data Modeling with Atmosphere, Polarization, and Chlorophyll Fluorescence. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2640–2649. [[CrossRef](#)]
28. Pinty, B.; Gobron, N.; Widlowski, J.-L.; Gerstl, S.A.W.; Verstraete, M.M.; Antunes, M.; Bacour, C.; Gascon, F.; Gastellu, J.-P.; Goel, N.; et al. Radiation transfer model intercomparison (RAMI) exercise. *J. Geophys. Res. Atmos.* **2001**, *106*, 11937–11956. [[CrossRef](#)]
29. Pinty, B.; Widlowski, J.L.; Taberner, M.; Gobron, N.; Verstraete, M.; Disney, M.; Gascon, F.; Gastellu-Etchegorry, J.-P.; Jiang, L.; Kuusk, A.; et al. Radiation Transfer Model Intercomparison (RAMI) exercise: Results from the second phase. *J. Geophys. Res.* **2004**, *109*, D06210. [[CrossRef](#)]
30. Widlowski, J.L.; Taberner, M.; Pinty, B.; Bruniquel, V.; Disney, M.; Fernandes, R.; Gastellu-Etchegorry, J.-P.; Gobron, N.; Kuusk, A.; Laverne, T.; et al. Third Radiation Transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance models. *J. Geophys. Res.* **2007**, *112*, D09111. [[CrossRef](#)]
31. Widlowski, J.L.; Pinty, B.; Lopatka, M.; Atzberger, C.; Buzica, D.; Chelle, M.; Disney, M.; Gastellu-Etchegorry, J.-P.; Gerboles, M.; Gobron, N.; et al. The 4th radiation transfer model intercomparison (RAMI-IV): Proficiency testing of canopy reflectance models with ISO-13528. *J. Geophys. Res.* **2013**, *118*, 6869–6890. [[CrossRef](#)]
32. Widlowski, J.L.; Robustelli, M.; Disney, M.; Gastellu-Etchegorry, J.-P.; Laverne, T.; Lewis, P.; North, P.R.J.; Pinty, B.; Thompson, R.; Verstraete, M. The RAMI On-line Model Checker (ROMC): A web-based benchmarking facility for canopy reflectance models. *Remote Sens. Environ.* **2007**, *112*, 1144–1150. [[CrossRef](#)]
33. Gastellu-Etchegorry, J.-P.; Guillevic, P.; Zagolski, F.; Demarez, V.; Trichon, V.; Deering, D.; Leroy, M. Modeling BRF and Radiation Regime of Boreal and Tropical Forests. *Remote Sens. Environ.* **1999**, *68*, 281–316. [[CrossRef](#)]

34. Zhen, Z.; Gastellu-Etchegorry, J.-P.; Chen, S.; Yin, T.; Chavanon, E.; Lauret, N.; Guilleux, J. Quantitative Analysis of DART Calibration Accuracy for Retrieving Spectral Signatures Over Urban Area. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 10057–10068. [\[CrossRef\]](#)
35. Gastellu-Etchegorry, J.-P.; Yin, T.; Lauret, N.; Cajgfinger, T.; Gregoire, T.; Grau, E.; Féret, J.-B.; Lopes, M.; Guilleux, J.; Dedieu, G.; et al. Discrete Anisotropic Radiative Transfer (DART 5) for Modeling Airborne and Satellite Spectroradiometer and LIDAR Acquisitions of Natural and Urban Landscapes. *Remote Sens.* **2015**, *7*, 1667–1701. [\[CrossRef\]](#)
36. Yan, G.; Jiao, Z.-H.; Wang, T.; Mu, X. Modeling surface longwave radiation over high-relief terrain. *Remote Sens. Environ.* **2020**, *237*, 111556. [\[CrossRef\]](#)
37. Dissegna, A.; Yin, T.; Wu, H.; Lauret, N.; Wei, S.; Gastellu-Etchegorry, J.-P.; Grêt-Regamey, A. Modeling Mean Radiant Temperature Distribution in Urban Landscapes Using DART. *Remote Sens.* **2021**, *13*, 1443. [\[CrossRef\]](#)
38. Wang, Y.; Lauret, N.; Gastellu-Etchegorry, J.-P. DART radiative transfer modelling for sloping landscapes. *Remote Sens. Environ.* **2020**, *247*, 111902. [\[CrossRef\]](#)
39. Zhang, Q.; Wang, D.; Gastellu-Etchegorry, J.-P.; Jinxin, Y.; Qian, Y. Impact of 3-D structures on directional effective emissivity in urban areas based on DART model. *Build. Environ.* **2023**, *239*, 110410. [\[CrossRef\]](#)
40. Stewart, I.D.; Oke, T. Local Climate Zones for Urban Temperature Studies. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 1879–1900. [\[CrossRef\]](#)
41. Gutman, G.; Ignatov, A. The Derivation of the Green Vegetation Fraction from NOAA/AVHRR Data for Use in Numerical Weather Prediction Models. *Int. J. Remote Sens.* **1998**, *19*, 1533–1543. [\[CrossRef\]](#)
42. Frederic, B.; Clevers, J.G.P.W.; Steven, M. The Robustness of Canopy Gap Fraction Estimates from Red and Near-Infrared Reflectances: A Comparison of Approaches. *Remote Sens. Environ.* **1995**, *54*, 141–151. [\[CrossRef\]](#)
43. Carlson, T.; Ripley, D. On the Relation between NDVI, Fractional Vegetation Cover, and Leaf Area Index. *Remote Sens. Environ.* **1997**, *62*, 241–252. [\[CrossRef\]](#)
44. Yang, J. Study of Urban Geometric Effects on Urban Surface Temperature Retrieval and Energy Flux Modelling: An Application in Hong Kong. Ph.D. Thesis, Hong Kong Polytechnic University, Hong Kong, China, 2017.
45. Kokalj, Ž.; Zaksek, K.; Oštir, K.; Pehani, P.; Čotar, K.; Somrak, M. Relief Visualization Toolbox, ver. 2.2.1 Manual. 2019. Available online: https://iaps.zrc-sazu.si/sites/default/files/rvt_2.2.1_0.pdf (accessed on 29 December 2024). [\[CrossRef\]](#)
46. Liu, X.; Chen, Y.; Mu, X.; Yan, G.; Xie, D.; Ma, X.; Yan, K.; Song, W.; Liu, Z. Correction for the Sun-Angle Effect on the NDVI Based on Path Length. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–17. [\[CrossRef\]](#)
47. Bhang, K.J.; Lee, J.-D. Consideration of NDVI and Surface Temperature Calculation from Satellite Imagery in Urban Areas: A Case Study for Gumi, Korea. *J. Korean Soc. Surv. Geod. Photogramm. Cartogr.* **2017**, *35*, 23–30. [\[CrossRef\]](#)
48. Jiang, Z.; Huete, A.; Chen, J.; Chen, Y.; Li, J.; Yan, G.; Zou, Y. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. *Remote Sens. Environ.* **2006**, *101*, 366–378. [\[CrossRef\]](#)
49. Peddle, D.; Brunke, S.; Hall, F.; Johnson, R. A Comparison of Spectral Mixture Analysis and Ten Vegetation Indices for Estimating Boreal Forest Biophysical Information from Airborne Data. *Can. J. Remote Sens.* **2001**, *27*, 627–635. [\[CrossRef\]](#)
50. Adeline, K.R.M.; Chen, M.; Briottet, X.; Pang, S.K.; Paparoditis, N. Shadow detection in very high spatial resolution aerial images: A comparative study. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 21–38. [\[CrossRef\]](#)
51. Thomas, C.; Doz, S.; Briottet, X.; Lachérade, S. AMARTIS v2: 3D Radiative Transfer Code in the [0.4; 2.5 μ m] Spectral Domain Dedicated to Urban Areas. *Remote Sens.* **2011**, *3*, 1914–1942. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.