First Experience with Zhuhai-1 Hyperspectral Data for Urban Dominant Tree Species Classification in Shenzhen, China

Haiming Qin 1, Weimin Wang 2,3,4, Yang Yao 1, Yuguo Qian 1, Xiangyun Xiong 2,3,4 and Weiqi Zhou 1,5,6,*©

1 State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China
2 Shenzhen Ecological and Environmental Monitoring Center of Guangdong Province, Shenzhen 518049, China
3 Guangdong Greater Bay Area, Change and Comprehensive Treatment of Regional Ecology and Environment, National Observation and Research Station, Shenzhen 518049, China
4 State Environmental Protection Scientific Observation and Research Station for Ecology and Environment of Rapid Urbanization Region, Shenzhen 518049, China
5 University of Chinese Academy of Sciences, Beijing 100049, China
6 Beijing-Tianjin-Hebei Urban Megaregion National Observation and Research Station for Eco-Environmental Change, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

* Correspondence: wzhou@rcees.ac.cn

Abstract: An accurate spatial distribution map of the urban dominant tree species is crucial for evaluating the ecosystem service value of urban forests and formulating urban sustainable development strategies. Spaceborne hyperspectral remote sensing has been utilized to distinguish tree species, but these hyperspectral data have a low spatial resolution (pixel size ≥ 30 m), which limits their ability to differentiate tree species in urban areas characterized by fragmented patches and robust spatial heterogeneity. Zhuhai-1 is a new hyperspectral satellite sensor with a higher spatial resolution of 10 m. This study aimed to evaluate the potential of Zhuhai-1 hyperspectral imagery for classifying the urban dominant tree species. We first extracted 32 reflectance bands and 18 vegetation indices from Zhuhai-1 hyperspectral data. We then used the random forest classifier to differentiate 28 dominant tree species in Shenzhen based on these hyperspectral features. Finally, we analyzed the effects of the classification paradigm, classifier, and species number on the classification accuracy. We found that combining the hyperspectral reflectance bands and vegetation indices could effectively distinguish the 28 dominant tree species in Shenzhen based on these hyperspectral features. Finally, we analyzed the effects of the classification paradigm, classifier, and species number on the classification accuracy. We found that combining the hyperspectral reflectance bands and vegetation indices could effectively distinguish the 28 dominant tree species in Shenzhen, obtaining an overall accuracy of 76.8%. Sensitivity analysis results indicated that the pixel-based classification paradigm was slightly superior to the object-based paradigm. The random forest classifier proved to be the optimal classifier for distinguishing tree species using Zhuhai-1 hyperspectral imagery. Moreover, reducing the species number could slowly improve the classification accuracy. These findings suggest that Zhuhai-1 hyperspectral data can identify the urban dominant tree species with accuracy and holds potential for application in other cities.

Keywords: Zhuhai-1; hyperspectral; tree species classification; urban; satellite

1. Introduction

Urban forests play a significant part in the urban ecosystem, supplying various ecosystem services such as fixing CO₂ and releasing O₂, lowering urban air temperatures, mitigating urban air pollution, isolating noise, alleviating urban floods, and providing habitats for animals [1–4]. However, there are significant differences in ecosystem service functions among different tree species [5]. The accurate evaluation of the ecosystem service effect of urban forests depends on accurate information regarding tree species composition and spatial distribution, which is also the premise of quantitative monitoring and effective evaluation of urban green quantity, biomass, and carbon stock [6–12]. In addition, tree
species diversity is a vital parameter to characterize urban ecosystems [13]. Accurately understanding the spatial patterns of the urban dominant tree species is crucial in better evaluating the ecosystem service value of urban dominant trees, improving urban environments, and formulating urban sustainable development strategies.

The traditional method of urban dominant tree species classification is to sample randomly in the city and then investigate the tree species. This method always takes a long time and has a high cost, making it hard to implement in large areas. Remote sensing can obtain the ground surface characteristics in large regions accurately and quickly, which provides a good opportunity for the timely and low-cost classification of the urban dominant tree species. Earlier research utilized multispectral remote sensing imagery such as Landsat, Sentinel-2, SPOT, WorldView, and Quickbird to distinguish the urban dominant tree species [14–18]. For example, Gavier-Pizarro et al. extracted the NDVI and the brightness, greenness, and wetness components of the Tasseled Cap Transformation from Landsat data to map the spatial pattern of glossy privet in an urban area of Argentina, obtaining an overall accuracy of 84% [19]. Poortinga et al. extracted a series of vegetation indices from Landsat and Sentinel-2 data to map rubber, palm oil, and mangrove, achieving an overall accuracy of 84% [20]. Pu et al. extracted a great deal of spectral and textural features from sunlit WorldView-2 imagery to identify seven urban tree species in the city of Tampa, FL, USA, using an object-based method, achieving an overall accuracy of 67.2% [18]. Nevertheless, because of inadequate spectral information in multispectral data, tree species classification accuracy is restricted.

Hyperspectral remote sensing imagery has notable advantages in classifying various species due to its fine spectral resolution, which enables the accurate detection of small spectral differences among various ground objects. Several studies have previously used airborne hyperspectral imagery to distinguish tree species, yielding high classification accuracy [21–24]. However, the cost of obtaining airborne hyperspectral imagery is relatively high, and airborne hyperspectral data can not be obtained in many areas because of the policy of banning flights in airspace. Meanwhile, several types of spaceborne hyperspectral data (for example, HJ-1A, Hyperion, GF-5, and PRISMA) have also been used to classify tree species [25–28]. However, these hyperspectral data have a low spatial resolution (pixel size ≥ 30 m) and are not suitable for tree species classification in urban regions, where green patches are fragmented and spatially heterogeneous.

Zhuhai-1 is a new hyperspectral satellite launched by China in 2018. It is equipped with an innovative hyperspectral sensor with 32 spectral bands, which is significantly more than those of multispectral sensors. The spatial resolution of Zhuhai-1 imagery is 10 m [29], finer than other hyperspectral satellite data (pixel size ≥ 30 m). The high spatial resolution combined with the high spectral bands provided by Zhuhai-1 are expected to offer useful information for ecology applications. Zhuhai-1 hyperspectral data has proved beneficial in various applications, including land cover classification, vegetation parameter estimation, and water quality parameter estimation with positive outcomes [30–32]. However, no attempt has been made to use Zhuhai-1 hyperspectral data to identify the urban dominant tree species.

In this research, the new hyperspectral satellite Zhuhai-1 data was used to differentiate the urban dominant tree species for the first time. Firstly, we carried out field investigation and identification of the dominant tree species. Next, we extracted hyperspectral features from Zhuhai-1 hyperspectral data. Then, we used these features to differentiate the dominant tree species in urban areas. Following that, we examined how the classification paradigm, classifier, and species number impacted the classification accuracy. We tested this approach in Shenzhen, Guangdong Province, China, where there were 28 dominant tree species. The specific purpose was to evaluate the effectiveness of the new hyperspectral satellite Zhuhai-1 in classifying the urban dominant tree species and to identify the optimal classification paradigm and classifier.
2. Materials and Methods

2.1. Study Region

The study region was situated in Shenzhen City, Guangdong Province, southern China (113°45′44″E–114°37′21″E, 22°26′59″N–22°51′49″N; Figure 1a–c) and belongs to the subtropical marine climate zone [33]. The area is about 1997 km$^2$, and the highest elevation is 943.7 m. The forest coverage of this study site was 55.56% in 2017 [34], and the forest is classified as a subtropical evergreen broadleaf forest.

Figure 1. The study area located in Shenzhen, southern China.

2.2. Zhuhai-1 Hyperspectral Remote Sensing Data

The Zhuhai-1 Orbita Hyper Spectral satellite equipped with a hyperspectral sensor was launched on 26 April 2018 by Zhuhai Orbita Aerospace Technology Co., Ltd. (Zhuhai, China). The satellite flies at 500 km altitude and captures hyperspectral images at a spatial resolution of 10 m. Zhuhai-1 data has 32 spectral bands; the detailed band information is shown in Table 1. In this research, we used four cloudless Zhuhai-1 hyperspectral images from October 1 and 16, 2019 to classify the urban dominant tree species. The weather was sunny and cloudless when the data were obtained.
Table 1. Detailed spectral information of Zhuhai-1 hyperspectral imagery.

<table>
<thead>
<tr>
<th>Band No.</th>
<th>Central Wavelength (nm)</th>
<th>Band No.</th>
<th>Central Wavelength (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>466</td>
<td>17</td>
<td>716</td>
</tr>
<tr>
<td>2</td>
<td>480</td>
<td>18</td>
<td>730</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>19</td>
<td>746</td>
</tr>
<tr>
<td>4</td>
<td>520</td>
<td>20</td>
<td>760</td>
</tr>
<tr>
<td>5</td>
<td>536</td>
<td>21</td>
<td>776</td>
</tr>
<tr>
<td>6</td>
<td>550</td>
<td>22</td>
<td>790</td>
</tr>
<tr>
<td>7</td>
<td>566</td>
<td>23</td>
<td>806</td>
</tr>
<tr>
<td>8</td>
<td>580</td>
<td>24</td>
<td>820</td>
</tr>
<tr>
<td>9</td>
<td>596</td>
<td>25</td>
<td>836</td>
</tr>
<tr>
<td>10</td>
<td>610</td>
<td>26</td>
<td>850</td>
</tr>
<tr>
<td>11</td>
<td>626</td>
<td>27</td>
<td>866</td>
</tr>
<tr>
<td>12</td>
<td>640</td>
<td>28</td>
<td>880</td>
</tr>
<tr>
<td>13</td>
<td>656</td>
<td>29</td>
<td>896</td>
</tr>
<tr>
<td>14</td>
<td>670</td>
<td>30</td>
<td>910</td>
</tr>
<tr>
<td>15</td>
<td>686</td>
<td>31</td>
<td>926</td>
</tr>
<tr>
<td>16</td>
<td>700</td>
<td>32</td>
<td>940</td>
</tr>
</tbody>
</table>

2.3. Land Cover Data

Qian et al. [34] obtained a land cover distribution map of Shenzhen in 2017 by utilizing high-resolution remote sensing data from SPOT 6 with a spatial resolution of 1.5 m. They classified the land cover types in Shenzhen into eight types, which were tree, grass, water, bare soil, building, road, construction, and impervious surface. The land cover type data had a classification accuracy of 88.67% with a corresponding kappa coefficient of 0.86 [34]. In this research, we classified tree species in the tree area of Shenzhen. Therefore, we extracted urban forest area from the land cover classification results of Qian et al. [34]. The forest area in Shenzhen is shown in Figure 1c.

2.4. Field Investigation Data

In order to collect field investigation data, we carried out field investigations in July 2018 and July 2019. We selected 4823 patches from the urban forest areas of the study region, and then the dominant tree species of each patch were visually identified in the field by plant ecologists. A high-precision Global Positioning System (GPS) was used to measure the geographical coordinates of each patch. The field investigation results showed that there were 28 main dominant tree species in Shenzhen, which were *Eucalyptus robusta*, *Litchi chinensis*, *Acacia mangium*, *Acacia confuse*, *Acacia auriculiformis*, *Acacia conferta*, *Dimocarpus longan*, *Ficus concinna*, *Cinnamomum camphora*, *Pinus massoniana*, *Schima superba*, *Sonneratia apetala*, *Delonix regia*, *Terminalia neotaliala*, *Roystonea regia*, *Ficus stipulosa*, *Bauhinia purpurea*, *Falcatoria falcata*, *Mangifera indica*, *Casuarina equisetifolia*, *Mimosa bimucronata*, *Leucaena leucocephala*, *Bombax ceiba*, *Ficus benjamina*, *Bischofia javanica*, *Alstonia scholaris*, *Khaya senegalensis*, and *Ficus altissima* from greatest to least. Figure 2 displays the spatial distribution of all field samples, covering a total of 101,384 pixels of the Zhuhai-1 hyperspectral imagery, which was about 10.14 km² in area. The average spectral reflectance for the 28 main tree species can be observed in Figure 3. The forested areas of the southern part had relatively few field samples because the forests in this area are continuously and widely distributed and densely grown, making it difficult for people to enter these areas for field investigation. Although the number of field samples in the southern forested areas was relatively small, the dominant tree species in these areas were relatively simple and basically consistent with the surrounding areas.
The dominant tree species in these areas were relatively simple and basically consistent with the surrounding areas.

Figure 2. Field samples used in this study.

Figure 3. Average reflectance according to spectral bands for 28 main tree species.

2.5. Methods

This study employed machine learning algorithms to classify the urban dominant tree species at both the pixel and object levels using Zhuhai-1 hyperspectral data. Figure 4 illustrates the process flowchart, which comprises five steps: (1) preprocessing raw hyperspectral data; (2) segmenting images; (3) extracting hyperspectral features; (4) classifying the urban dominant tree species using machine learning methods; and (5) evaluation of
classification accuracy. Finally, the optimal classification model was employed to produce a spatial distribution map of the dominant tree species.

2.5.1. Data Preprocessing

To obtain the hyperspectral reflectance bands, we preprocessed the Zhuhai-1 hyperspectral remote sensing data using ENVI 5.3 software. At first, we transformed the Zhuhai-1 hyperspectral digital value into radiance at the atmosphere’s top using the standard radiometric calibration formula and coefficients provided by the data producer [35]. Next, we used the FLAASH atmospheric correction module embedded in ENVI 5.3 software, in which the tropical atmospheric model and the urban aerosol model were the key input parameters, to carry out atmospheric correction of the hyperspectral data to obtain the surface reflectance [36]. Then, we used the seamless mosaic tool to mosaic the four hyperspectral images that covered the study area. Finally, we used a polygon of Shenzhen City as the mask to extract the hyperspectral data. After the above data preprocessing, we obtained 32 reflectance bands from the Zhuhai-1 hyperspectral data in the entire study region.

2.5.2. Image Segmentation

We used two classification paradigms to classify tree species: pixel- and object-based paradigms. For the object-based tree species classification, image segmentation was first carried out on the hyperspectral data. We used the multiresolution segmentation method embedded in Trimble eCognition Developer 9.5.1 to produce objects from the Zhuhai-1 hyperspectral images [37]. Multiresolution segmentation method refers to the generation of image objects with minimum heterogeneity and maximum homogeneity at any scale on the premise of minimizing the loss of image information [37,38]. Its key parameters

Figure 4. The flowchart for urban dominant tree species classification using Zhuhai-1 hyperspectral data.

- Zhuhai-1 hyperspectral data
- Field sample data
- Data pre-processing
  - Pixel-based paradigm
  - Object-based paradigm
  - Image segmentation
  - Feature extraction
    - 32 band reflectance
    - 18 vegetation indices
  - Urban tree species classification using machine learning algorithms
  - Training sample
  - Urban tree species distribution map
  - Testing sample
  - Accuracy assessment

- Figure 4 illustrates the flowchart for urban dominant tree species classification using Zhuhai-1 hyperspectral data.
include the scale and compactness/smoothness weight [37]. Thirty-two bands of Zhuhai-1 hyperspectral data were used as input. We tested the parameters based on experience and visually evaluated the segmentation results. As a result, we set the scale as 10 and the compactness/smoothness weight as 0.5/0.5.

2.5.3. Hyperspectral Feature Extraction

For each pixel and object, 50 features were extracted from the Zhuhai-1 hyperspectral data, including 32 spectral reflectance bands and 18 vegetation indices. At the pixel level, the 32 spectral reflectance bands were obtained directly from the preprocessed hyperspectral image. The 18 vegetation indices of each pixel were calculated using ENVI 5.3 software [33]. These vegetation indices can characterize the structural and physiological features of trees and have demonstrated effectiveness in the classification of tree species [39–41]. Table 2 lists the vegetation indices and their formulas used in this study. At the object level, 32 reflectance bands and 18 vegetation indices were determined according to the average value of all pixels in each object.

Table 2. Hyperspectral vegetation indices and their formulas.

<table>
<thead>
<tr>
<th>Information Types</th>
<th>Metrics</th>
<th>Formula</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf area and canopy structure</td>
<td>Normalized Difference Vegetation Index (NDVI)</td>
<td>( \text{NDVI} = \frac{\rho_{790} - \rho_{670}}{\rho_{790} + \rho_{670}} )</td>
<td>[42]</td>
</tr>
<tr>
<td>Soil Adjusted Vegetation Index</td>
<td>Soil Adjusted Vegetation Index [43]</td>
<td>( \text{SAVI} = \frac{1.5 \times (\rho_{790} - \rho_{670})}{\rho_{790} + \rho_{670} + 0.5} )</td>
<td>[44]</td>
</tr>
<tr>
<td>Index (ARVI)</td>
<td>Atmospherically Resistant Vegetation Index (ARVI)</td>
<td>( \text{ARVI} = \frac{\rho_{790} - 2 \times \rho_{670} + \rho_{480}}{\rho_{790} + 2 \times \rho_{670} - \rho_{480}} )</td>
<td>[45]</td>
</tr>
<tr>
<td>Enhanced Vegetation Index (EVI)</td>
<td>Enhanced Vegetation Index (EVI)</td>
<td>( \text{EVI} = 2.5 \times \frac{1 + \rho_{686} + 6 \times \rho_{670} - 7.5 \times \rho_{480}}{\rho_{790}} )</td>
<td>[46]</td>
</tr>
<tr>
<td>Modified Red Edge Normalized</td>
<td>Modified Red Edge Normalized Difference Vegetation Index (MRENDVI)</td>
<td>( \text{MRENDVI} = \frac{\rho_{746} - \rho_{700}}{\rho_{746} + \rho_{700} - 2 \times \rho_{466} - 1} )</td>
<td>[47]</td>
</tr>
<tr>
<td>Difference Vegetation Index (MRESRI)</td>
<td>Modified Red Edge Simple Ratio Index (MRESRI)</td>
<td>( \text{MRESRI} = \frac{\rho_{716} - \rho_{466}}{\rho_{716} + \rho_{466}} )</td>
<td>[48]</td>
</tr>
<tr>
<td>Vogelmann Red Edge Index (VOG)</td>
<td>Vogelmann Red Edge Index (VOG)</td>
<td>( \text{VOG} = \frac{\rho_{716}}{\rho_{746}^{\text{746}} - \rho_{686}} )</td>
<td>[49]</td>
</tr>
<tr>
<td>Mean Value of Red Edge (Mean (686-749))</td>
<td>Mean Value of Red Edge (Mean (686-749))</td>
<td>( \text{Mean}<em>{686-749} = \frac{\sum</em>{i=746}^{i=686} \rho_{i}}{n} )</td>
<td>[46,50]</td>
</tr>
<tr>
<td>Slope Location of Red Edge (SL)</td>
<td>Slope Location of Red Edge (SL)</td>
<td>( \text{SL} = \frac{\rho_{746} - \rho_{686}}{60} )</td>
<td>[50]</td>
</tr>
<tr>
<td>Leaf and canopy pigments</td>
<td>Datt Chlorophyll Content Index (Datt)</td>
<td>( \text{Datt} = \frac{\rho_{850} - \rho_{716}}{\rho_{580} - \rho_{716}} )</td>
<td>[51]</td>
</tr>
<tr>
<td>Chlorophyll Index (CI)</td>
<td>Chlorophyll Index (CI)</td>
<td>( \text{CI} = \frac{\rho_{850} - \rho_{716}}{\rho_{750} - \rho_{716}} - 1 )</td>
<td>[52]</td>
</tr>
<tr>
<td>Red Edge Index (REI)</td>
<td>Red Edge Index (REI)</td>
<td>( \text{REI} = \frac{\rho_{716}}{\rho_{746}} - 1 )</td>
<td>[53]</td>
</tr>
<tr>
<td>Green Index (GI)</td>
<td>Green Index (GI)</td>
<td>( \text{GI} = \frac{\rho_{806}}{\rho_{750}} - 1 )</td>
<td>[54]</td>
</tr>
<tr>
<td>Plant stress</td>
<td>Plant Stress Index (PSI)</td>
<td>( \text{PSI} = \frac{\rho_{700}}{\rho_{750} - \rho_{590}} )</td>
<td>[55]</td>
</tr>
<tr>
<td>Ratio Index (RI)</td>
<td>Ratio Index (RI)</td>
<td>( \text{RI} = \frac{\rho_{700}}{\rho_{750}} )</td>
<td>[55]</td>
</tr>
<tr>
<td>Red Edge Vegetation Pressure Index</td>
<td>Red Edge Vegetation Pressure Index (RVSI)</td>
<td>( \text{RVSI} = \frac{\rho_{716} + \rho_{746}}{2 - \rho_{730}} )</td>
<td>[56]</td>
</tr>
<tr>
<td>Light energy utilization efficiency</td>
<td>Structure Insensitive Pigment Index [57]</td>
<td>( \text{SIPI} = \frac{\rho_{806} - \rho_{466}}{\rho_{806} + \rho_{866}} )</td>
<td>[54]</td>
</tr>
<tr>
<td>Modified Photochemical Reflectance Index (MPRI)</td>
<td>Modified Photochemical Reflectance Index (MPRI)</td>
<td>( \text{MPRI} = \frac{\rho_{520} - \rho_{536}}{\rho_{520} + \rho_{536}} )</td>
<td>[58]</td>
</tr>
</tbody>
</table>
2.5.4. Tree Species Classification

A random forest classifier was used to classify the urban dominant tree species in this study. The random forest classifier was proposed by Leo Breiman [59]. It contains multiple decision trees, and the mode of categories obtained by all the decision trees determines the category of the classification. The random forest classifier has strong robustness and can deal with multiple collinear relationships well. Two parameters (ntree and mtry) were necessary for modeling; these were set to 500 and 7, respectively, according to Immitzer et al. [17,60] Because of its high classification accuracy, the random forest classifier has been successfully used in tree species classification [26,61–64]. In this study, we input 32 hyperspectral reflectance bands and 18 vegetation indices into the random forest classifier to classify the urban dominant tree species and identify important features. We randomly selected 70% of the field samples of each tree species to train the random forest classifier and then used the remaining 30% of the field samples for verification purposes.

2.5.5. Accuracy Assessment

To evaluate the accuracy of tree species classification, we utilized the confusion matrix method (Table 3), which is a widely accepted technique in assessing classification accuracy [33,65,66]. The accuracy metrics include user accuracy, producer accuracy, overall accuracy, and the kappa coefficient [67]. The calculation methods of producer accuracy and user accuracy are shown in Table 3. The overall accuracy is determined by the proportion of correctly classified pixels or objects to the total number of pixels or objects; its calculation method is shown in Equation (1). The kappa coefficient is a statistical measure of the agreement or reliability of the classification outcomes; its calculation is shown in Equations (2) and (3). For specifics regarding the calculation approaches for user accuracy, producer accuracy, overall accuracy, and the kappa coefficient, please refer to Qin et al. [33,68].

\[
OA = \frac{a + c + i}{r} \tag{1}
\]

\[
p_e = \frac{o \times l + p \times m + q \times n}{r \times r} \tag{2}
\]

\[
k = \frac{OA - p_e}{1 - p_e} \tag{3}
\]

Table 3. An example of the confusion matrix of the tree species classification results.

<table>
<thead>
<tr>
<th>Observed Types</th>
<th>Predicted Types</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>Sum</th>
<th>Producer Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>1</td>
<td></td>
<td>7/7</td>
</tr>
<tr>
<td>B</td>
<td>d</td>
<td>e</td>
<td>f</td>
<td>m</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>g</td>
<td>h</td>
<td>i</td>
<td>n</td>
<td>7/7</td>
<td></td>
</tr>
<tr>
<td>Sum</td>
<td>o</td>
<td>p</td>
<td>q</td>
<td>r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>User accuracy</td>
<td>7/7</td>
<td>7/7</td>
<td>7/7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Results

3.1. Feature Importance

Figure 5 displays the variable importance of all 50 hyperspectral features for classifying the urban dominant tree species using the random forest classifier. The three most important features were hyperspectral reflectance, which were b1, b2, and b11, respectively. Among the hyperspectral vegetation indices, MRESRI held the most significance, followed by MRENDVI and GI. The total percentage of variable importance for the first 18 features was over 50%, and that of the first 30 features was over 70%. Among the top 18 important features, 7 features were hyperspectral reflectance bands and 11 features were hyperspectral vegetation indices. Among the 30 most important features, 14 features were
hyperspectral vegetation indices and 16 features were hyperspectral reflectance. In general, the contributions of the hyperspectral reflectance bands and vegetation indices to the tree species classification were similar.

**Figure 5.** The variable importance of all hyperspectral features for classification of tree species using the random forest classifier.

### 3.2. Pixel-Based Tree Species Classification Using the RF Classifier

Table 4 shows the classification accuracy of 28 main species based on three feature sets, namely the hyperspectral reflectance bands, the vegetation indices, and a combination of both. All three feature sets resulted in satisfactory classification results. The classification accuracy of the 32 hyperspectral reflectance bands (OA = 76.5% and kappa = 0.75) was slightly higher than that of the 18 vegetation indices (OA = 75.6% and kappa = 0.74). Combining the reflectance bands with vegetation indices could marginally enhance the classification accuracy (overall accuracy = 76.8% and kappa = 0.75). The pixel-based tree species classification results using the random forest classifier derived from all hyperspectral features are presented in Figure 6a. Although the user accuracy across all tree species was well-balanced, the producer accuracy was largely inconsistent. All results from the classification accuracy assessment using all hyperspectral features are shown in Table 5. The most accurately classified tree species were *Schima superba, Sonneratia apetala, Terminalia neotaliala, Mangifera indica, Bischofia javanica, Alstonia scholaris*, and *Khaya senegalensis*, with producer and user accuracies exceeding 80%.

**Table 4.** The classification results of 28 tree species in Shenzhen based on three feature sets.

<table>
<thead>
<tr>
<th>Features</th>
<th>Overall Accuracy</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>32 reflectance bands</td>
<td>76.5%</td>
<td>0.75</td>
</tr>
<tr>
<td>18 vegetation indices</td>
<td>75.6%</td>
<td>0.74</td>
</tr>
<tr>
<td>32 reflectance bands + 18 vegetation indices</td>
<td>76.8%</td>
<td>0.75</td>
</tr>
</tbody>
</table>
Figure 6. (a) Pixel-based tree species classification results for the entire study area derived from all hyperspectral features using the RF classifier; (b) detailed pixel-based tree species classification results for the area in the black box in (a); (c) detailed object-based tree species classification results for the area in the black box in (a).
Table 5. Classification accuracy of 28 tree species using the RF classifier based on 50 hyperspectral metrics.

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Producer Accuracy</th>
<th>User Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eucalyptus robusta</td>
<td>74.3%</td>
<td>68.0%</td>
</tr>
<tr>
<td>Litchi chinensis</td>
<td>82.4%</td>
<td>68.8%</td>
</tr>
<tr>
<td>Acacia mangium</td>
<td>77.2%</td>
<td>78.8%</td>
</tr>
<tr>
<td>Acacia confusa</td>
<td>70.0%</td>
<td>85.5%</td>
</tr>
<tr>
<td>Acacia auriculiformis</td>
<td>93.2%</td>
<td>70.3%</td>
</tr>
<tr>
<td>Acacia conferta</td>
<td>74.3%</td>
<td>82.9%</td>
</tr>
<tr>
<td>Dimocarpus longan</td>
<td>74.7%</td>
<td>72.0%</td>
</tr>
<tr>
<td>Ficus concinna</td>
<td>71.4%</td>
<td>77.5%</td>
</tr>
<tr>
<td>Cinnamomum camphora</td>
<td>63.4%</td>
<td>89.9%</td>
</tr>
<tr>
<td>Pinus massoniana</td>
<td>87.8%</td>
<td>76.6%</td>
</tr>
<tr>
<td>Schima superba</td>
<td>93.3%</td>
<td>97.9%</td>
</tr>
<tr>
<td>Sonneratia apetala</td>
<td>96.2%</td>
<td>94.7%</td>
</tr>
<tr>
<td>Delonix regia</td>
<td>92.6%</td>
<td>75.8%</td>
</tr>
<tr>
<td>Terminalia neotaliala</td>
<td>84.5%</td>
<td>86.3%</td>
</tr>
<tr>
<td>Roystonea regia</td>
<td>72.9%</td>
<td>66.2%</td>
</tr>
<tr>
<td>Ficus stipulosa</td>
<td>79.5%</td>
<td>67.8%</td>
</tr>
<tr>
<td>Bauhinia purpurea</td>
<td>91.4%</td>
<td>74.7%</td>
</tr>
<tr>
<td>Falcatafia falcata</td>
<td>2.6%</td>
<td>83.3%</td>
</tr>
<tr>
<td>Mangifera indica</td>
<td>92.9%</td>
<td>100%</td>
</tr>
<tr>
<td>Casuarina equisetifolia</td>
<td>79.5%</td>
<td>70.5%</td>
</tr>
<tr>
<td>Mimosa bimucronata</td>
<td>48.0%</td>
<td>92.3%</td>
</tr>
<tr>
<td>Leucaena leucocephala</td>
<td>73.2%</td>
<td>74.5%</td>
</tr>
<tr>
<td>Bombax ceiba</td>
<td>50.0%</td>
<td>100%</td>
</tr>
<tr>
<td>Ficus benjamina</td>
<td>73.3%</td>
<td>76.8%</td>
</tr>
<tr>
<td>Bischofia javanica</td>
<td>100%</td>
<td>84.2%</td>
</tr>
<tr>
<td>Alstonia scholaris</td>
<td>84.0%</td>
<td>80.8%</td>
</tr>
<tr>
<td>Khaya senegalensis</td>
<td>90.0%</td>
<td>100%</td>
</tr>
<tr>
<td>Ficus altissima</td>
<td>73.7%</td>
<td>93.3%</td>
</tr>
</tbody>
</table>

Overall Accuracy 76.8%  
Kappa 0.75

3.3. Pixel- vs. Object-Based Classification Results

We implemented object-based classification by utilizing the random forest classifier and all hyperspectral features to investigate whether it could improve the tree species classification accuracy. We found that the pixel-based paradigm (OA = 76.8% and kappa 0.75) was slightly superior to the object-based paradigm (OA = 76.5% and kappa = 0.74). Therefore, when using Zhuhai-1 hyperspectral data to distinguish the urban dominant tree species, the pixel-based classification paradigm was slightly better than object-based classification paradigm. Since there were no obvious differences between our pixel- and object-based classification results, only the pixel-based tree species classification results for the entire study area are displayed in Figure 6a. To show the detailed differences between the two results, we chose a small area of the study region (2 km × 2 km) to compare the pixel-based classification results (Figure 6b) and object-based classification results (Figure 6c), and the results showed that the main difference was that pixel-based classification result had more noise points.

3.4. Performance Comparison among the Four Classifiers

In order to explore whether different classifiers affected the accuracy of tree species differentiation, we employed the random forest, support vector machine, Bayes, and K-nearest neighbor classifiers to classify tree species based on all hyperspectral features. The results indicated that the RF classifier achieved the highest classification accuracy, followed by KNN (OA = 62.3% and kappa = 0.60), Bayes (OA = 61.4% and kappa = 0.59), and SVM (OA = 43.2% and kappa = 0.38) when distinguishing the urban dominant tree species using
Zhuhai-1 hyperspectral data. Therefore, the RF classifier was the recommended classifier for urban dominant tree species classification.

3.5. Effect of Species Number on Tree Species Classification

To investigate the impact of species number on tree species classification accuracy, we evaluated the classification accuracy using all hyperspectral features with the species number ranging from 5 to 28. We selected 5, 10, 15, 20, 25, and 28 tree species for classification according to the list of tree species sorted from greatest to least obtained in the field investigation (Section 2.4). We used the training samples of the selected tree species for classification, and then used the verification samples of these tree species to evaluate the accuracy. Therefore, we did not delete the pixels dominated by the eliminated species during the classification process. The relationship between the classification accuracy and species number is illustrated in Figure 7. We found that as the species number increased from 5 to 28, the classification accuracy decreased slowly. However, when the species number increased from 15 to 20, the classification accuracy decreased obviously. The overall accuracy was 82.6% when there were five tree species, while it decreased to only 76.8% when there were 28 tree species.

![Figure 7. The relationship between the species number and classification accuracy.](image)

4. Discussion

4.1. Performance of Zhuhai-1 Hyperspectral Data in Urban Dominant Tree Species Classification

This research first assessed the capability of the new hyperspectral satellite Zhuhai-1 imagery to differentiate the urban dominant tree species, achieving a satisfactory classification accuracy of 76.8%. It outperformed the previous studies employing multispectral satellite data such as that from Sentinel-2, Landsat, and Ziyuan-3 [17,66,69–71]. The higher spectral resolution and additional spectral bands of the Zhuhai-1 imagery make it more suitable for distinguishing tree species in regions with abundant species. The higher spatial resolution of the Zhuhai-1 hyperspectral data compared to other hyperspectral satellite data such as HyspIRI (30 m), Hyperion (30 m), HJ-1A (100 m), PRISMA (30 m) returns better results in accurately classifying tree species in heterogeneous and fragmented regions [25,26,28,72]. The classification result was worse than in several existing studies [27,73–76]. The main reason was that there were more tree species in this research, which aggravated the difficulty of tree species differentiation. This was verified in our sensitivity analysis of the effect of the species number on classification accuracy (Section 3.5).
In addition, several studies used airborne or UAV hyperspectral data to classify tree species and obtained a higher classification accuracy than this research [23,33]. This was mainly due to the fact that airborne or UAV hyperspectral data have a higher spatial resolution and spectral resolution that are more capable of distinguishing tree species.

By comparing the performance of the Zhuhai-1 hyperspectral reflectance bands and vegetation indices in tree species differentiation, we found that there were no obvious differences between them, and the former was slightly better than the latter. This means that the reflectance bands of the Zhuhai-1 hyperspectral data contained enough information to distinguish the urban dominant tree species. However, several studies found that vegetation indices performed better than reflectance bands in distinguishing tree species [37,77]. This may have been due to the fact that the remote sensing data they used contained limited spectral bands (less than 10 bands), which were not enough to distinguish diverse tree species, so it was necessary to extract vegetation index that combined information from multiple bands to provide supplementary information for tree species classification. It may also have been due to the different spatial resolutions of the remote sensing images used in the different studies.

4.2. Effect of Different Classification Paradigms on Classification Accuracy

By comparing the classification accuracies obtained by the object- and pixel-based methods, we observed that their classification results had a subtle difference, and the pixel-based method was slightly better. Although several existing studies align with our findings [17,78], more studies have reported that the object-based method was more effective than the pixel-based method in classification [60,74,79]. This was because these studies used images with a very high spatial resolution (with pixel sizes of about 1 m) to distinguish tree species in which each tree crown covered many pixels. In these cases, the object-based method could better express the crown characteristics and obtain better tree species classification results. In this research, the Zhuhai-1 hyperspectral data had a spatial resolution equivalent to 10 m, which is quite similar to the size of a tree crown. Hence, the pixel-based method had the ability to deliver a more precise description of crown features when compared to the object-based method. Consequently, when utilizing the Zhuhai-1 hyperspectral data, the pixel-based method was better suited for the differentiation of the urban dominant tree species.

4.3. Effect of Different Classifiers on Classification Accuracy

In this study, we compared four machine learning methods; i.e., SVM, RF, k-NN, and Bayes, in classifying the urban dominant tree species. Our analysis revealed that the RF outperformed the other three classification methods. The results were in agreement with those found by earlier researchers [27,61,80,81]. The RF classifier can balance errors in the case of unbalanced data sets, so it can obtain better classification results than other classifiers. In urban areas, there are a variety of tree species and significant quantitative differences among species, which lead to an imbalance in the training sample sets. Therefore, the RF classifier is more suitable to distinguish the urban dominant tree species when using Zhuhai-1 hyperspectral data. Several studies reported that SVM performed better than RF in classification, which may have been due to the small number of samples [73,82].

4.4. Limitations and Future Research Prospects

There are obvious differences in ecological service function among different tree species. Based on the spatial distribution map of the urban dominant tree species obtained in this study, the ecological services of urban dominant trees can be evaluated more accurately. The urban dominant tree species classification results obtained by this study were satisfactory but not extremely high. One of the main limitations was the spectral similarity of the 28 dominant tree species, which posed a great challenge in the classification of tree species. In addition, although the spatial resolution of the Zhuhai-1 hyperspectral data is significantly finer than that of Landsat-like sensors, the presence of mixture pixels
introduced challenges to the classification of tree species, and this needs to be solved
in the future. Furthermore, all of the tree species in this study region were subtropical
broad-leaved forests with luxuriant tree growth and high morphological similarity, which
increased the difficulty of tree species classification.

Further efforts should be made to enhance the classification accuracy. Diverse tree
species exhibit notable disparities in their phenological characteristics, which subsequently
enhances their spectral distinctiveness. The potential of multitemporal Zhuhai-1 hyperspec-
tral data covering the key phenological periods in urban dominant tree species mapping
needs to be further explored. In addition to the spectral information provided by Zhuhai-1
hyperspectral data, it is imperative to examine whether incorporating topography, climate,
and vertical structure information can enhance the accuracy of tree species classification.
More artificial intelligence techniques should also be considered for fine tree species classi-
fication. In addition, it is also vital to verify the effectiveness of Zhuhai-1 hyperspectral
data in urban dominant tree species classification across more extensive regions.

5. Conclusions

This research assessed the effectiveness of data from the new hyperspectral satellite
Zhuhai-1 in mapping the urban dominant tree species in Shenzhen, southern China. We
extracted 32 reflectance bands and 18 vegetation indices from the Zhuhai-1 hyperspectral
data and then input them into four classifiers to distinguish 28 dominant tree species in
urban areas at the pixel and object levels. The results demonstrated that the Zhuhai-1
hyperspectral data could effectively distinguish the 28 dominant tree species in urban areas,
obtaining an OA of 76.8% and a kappa coefficient of 0.75. The hyperspectral reflectance
bands and vegetation indices contributed to the tree species classification similarly. The sen-
sitivity analysis results suggested that the pixel-based approach marginally outperformed
the object-based approach in classifying the urban dominant tree species using the Zhuhai-1
hyperspectral data, and the RF classifier demonstrated the best results among all classifiers
tested. Moreover, reducing the species number could improve the classification accuracy.
These findings provide a framework for urban dominant tree species classification using
Zhuhai-1 hyperspectral imagery, and the urban dominant tree species map generated in
this study holds potential for a wide range of practical applications.

Author Contributions: Conceptualization, H.Q. and W.Z.; methodology, H.Q., Y.Y. and Y.Q.; software,
writing—original draft preparation, H.Q. and Y.Y.; writing—review and editing, W.W., X.X. and W.Z.;
funding acquisition, H.Q. and W.Z. All authors have read and agreed to the published version of
the manuscript.

Funding: This research was funded by the National Key Research and Development Program of
China (grant number 2022YFF1301101); the National Natural Science Foundation of China (grant
numbers 32101292); the Research Center for Eco-Environmental Sciences, Chinese Academy of
Sciences (grant number RCEES-TDZ-2021-9); and the Shenzhen Ecological and Environmental Moni-
toring Center of Guangdong Province (grant numbers SZDL2023000962, GXZX-20210521SZGK, and
0733-22201076).

Data Availability Statement: Not applicable.

Acknowledgments: Not applicable.

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

References
1. Moody, R.; Geron, N.; Healy, M.; Rogan, J.; Martin, D. Modeling the spatial distribution of the current and future ecosystem
services of urban tree planting in Chicopee and Fall River, Massachusetts. Urban For. Urban Green. 2021, 66, 127403. [CrossRef]
3. Palliwoda, J.; Banzhaf, E.; Priess, J.A. How do the green components of urban green infrastructure influence the use of ecosystem
services? Examples from Leipzig, Germany. Landsc. Ecol. 2020, 35, 1127–1142. [CrossRef]


16. Lim, J.; Kim, K.-M.; Jin, R. Tree Species Classification Using Hyperion and Sentinel-2 Data with Machine Learning in South Korea. *ISPRS Int. J. Geo-Inf.* 2019, 8, 150. [CrossRef]

17. Wang, X.; Ren, H. DBMF: A Novel Method for Tree Species Fusion Classification Based on Multi-Source Images. *Forests* 2021, 13, 33. [CrossRef]


53. Raper, T.B. Effectiveness of Crop Reflectance Sensors on Detection of Cotton (Gossypium hirsutum L.) Growth and Nitrogen Status; Mississippi State University: Lee MS, USA, 2011.


61. Varin, M.; Chalghaf, B.; Joannis, G. Object-Based Approach Using Very High Spatial Resolution 16-Band WorldView-3 and LiDAR Data for Tree Species Classification in a Broadleaf Forest in Quebec, Canada. *Remote Sens.* 2020, 12, 3092. [CrossRef]
69. Dong, C.; Zhao, G.; Meng, Y.; Li, B.; Peng, B. The Effect of Topographic Correction on Forest Tree Species Classification Accuracy. *Remote Sens.* 2020, 12, 787. [CrossRef]