Land Cover and Forest Type Classification by Values of Vegetation Indices and Forest Structure of Tropical Lowland Forests in Central Vietnam

Hung Nguyen Trong, The Dung Nguyen, and Martin Kappas

1Department of Cartography, GIS and Remote Sensing, Georg-August-Universität Göttingen, 37077 Göttingen, Germany
2Ministry of Natural Resources and Environment, Hanoi, Vietnam
3Vietnam National University of Forestry, Xuan Mai, Hanoi, Vietnam

Correspondence should be addressed to Hung Nguyen Trong; tronghung2128@gmail.com

Received 16 July 2020; Revised 1 September 2020; Accepted 2 September 2020; Published 22 September 2020

Academic Editor: Anna Zróbek-Sokolnik

Copyright © 2020 Hung Nguyen Trong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims to (i) optimize the application of multiple bands of satellite images for land cover classification by using random forest algorithms and (ii) assess correlations and regression of vegetation indices of a better-performed land cover classification image with vertical and horizontal structures of tropical lowland forests in Central Vietnam. In this study, we used Sentinel-2 and Landsat-8 to classify seven land cover classes of which three forest types were substratified as undisturbed, low disturbed, and disturbed forests where forest inventory of 90 plots, as ground-truth, was randomly sampled to measure forest tree parameters. A total of 3226 training points were sampled on seven land cover types. The performance of Landsat-8 showed out-of-bag error of 31.6%, overall accuracy of 68%, kappa of 67.5%, while Sentinel-2 showed out-of-bag error of 14.3% and overall accuracy of 85.7% and kappa of 83%. Ten vegetation indices of the better-performed image were extracted to find out (i) the correlation and regression of horizontal and vertical structures of trees and (ii) assess the variation values between ground-truthing plots and training sample plots in three forest types. The result of the t test on vegetation indices showed that six out of ten vegetation indices were significant at \( p < 0.05 \). Seven vegetation indices had a correlation with the horizontal structure, but four vegetation indices, namely, Enhanced Vegetation Index, Perpendicular Vegetation Index, Difference Vegetation Index, and Transformed Normalized Difference Vegetation Index, had better correlations \( r = 0.66, 0.65, 0.65, 0.63 \) and regression results were of \( R^2 = 0.44, 0.43, 0.43, 0.40 \), respectively. The correlations of tree height were \( r = 0.46, 0.43, 0.43, \) and 0.49 and its regressions were of \( R^2 = 0.21, 0.19, 0.18, \) and 0.24, respectively. The results show the possibility of using random forest algorithm with Sentinel-2 in forest type classification in line with vegetation indices application.

1. Introduction

Forests, at a nationwide scale, need a monitoring system as fundamental tools to support the management of landscapes, land use, ecosystem, and biodiversity for multiple production purposes including national forest inventory (NFI) and international conventions [1–4]. It is clearly stated that land cover and land use mapping classification, derived from remote sensing data, for natural resource management, monitoring, and development strategies are still open as big demand of the society and can be expressed as discrete classes or continuous land cover attributes [5]. Many studies have made land cover maps from different data sources such as multispectral, hyperspectral, and radar aperture [6–9] and effortted to generating information from remote sensing (RS) by using machine learning algorithms to classify land cover [10–12]. Machine learning algorithms such as support vector machine, k-nearest neighbor, and random forest are non-parametric classifiers caused huge attention for remote sensing in the last decades. The random forest (RF) is one of the most optimal artificial methodologies used for land cover classification with nonparametric classification algorithms.
that can run on large data sets [13, 14]. Furthermore, RF is an optimal algorithm which is suitable for an excellent number of remote sensing applications and compared with other conventional techniques [15–17]. Remote sensing images provide potential information on tropical landscape forests and land use types [18] where the landscape is defined as a heterogeneous land area from a set of a cluster of interacting ecosystems that repeat in similar shape throughout and as an area that is spatially heterogeneous in at least one factor of interests [19, 20]. The structure of forest has a relation to spatial distribution and as an important factor of forest ecological processes which supports to give more patterns of some taxa and even the disturbance status [21, 22]. Forest structure provides important additional information to improve the estimation of forest stand variables [23]. High-resolution remote sensing images with the support of promising technique has potential for capturing the status but may not obtain forest productivity, which can partially reflect forest structure [23]. The outstanding method for evaluating land cover classification is a ground-based assessment derived from remote sensing [24]. Different sensors and methodology may give different results of application and information [25–27]. Vegetation indices (VIs) are essential for vegetation cover classification captured from the radiometric biophysical derivation and vegetation structure. These indices contribute to land use planning and manage natural resources management and provide to policy making [28–31]. Sentinel-2 and Landsat-8 were developed to support vegetation, land cover, land use, environmental monitoring, and others such as biophysical and geographical resources [32, 33]. The Sentinel-2 measures reflected radiance within 13 spectral bands, whereas Landsat-8 has eleven bands. Multi-spectral bands help to map the vegetation types of the regional scale [34, 35]. Some studies have tried to classify land cover and vegetation cover in tropical life form using vegetation indices and texture, but mapping different land cover and forest types in the tropic is a big challenge due to its heterogeneity of landscape and availability of optical satellite images with low cloud ratio [36, 37]. The spatial resolution is dependent on the particular spectral bands, but 10 m resolution of Sentinel-2 can provide feasible phenological values for different landscapes [38]. Several studies applied multispectral bands of Landsat-8 with a resolution of 30 m to assess vegetation dynamic [39]. The key process of the research workflow is presented in Figure 1.

No studies have been conducted applying RF for land cover classification in combination with extracted vegetation indices to stratify forest structures and to compare the values of vegetation indices between the ground-truth and training sample plots of VIs from the best-performed satellite images in the tropical lowland forests in Vietnam. This paper aims to (i) compare the application of multiple band Sentinel-2 and Landsat-8 images for land cover classification by applying RF and (ii) evaluate the correlation and regression of vegetation indices forest vertical and horizontal structures with the abovementioned two sensors in the tropical lowland forests in Central Vietnam.

2. Study Area

The study area is located between 107°.22’E to 107.30’E and 16°.02’N to 16°10’N in the Western part of Thua Thien Hue province (Figure 2). The average annual temperature is 21.9°C, the average high temperature is 25.3°C, and the average low temperature is 17°C. Moisture content varies from 81% to 94%, and it is normally influenced by an intertropical convergence zone that typically causes tropical low pressures, monsoons, and typhoons leading to annual rainfall of about 3500 mm with an average of 200 rainy days annually, most of which accumulates between September and December [40, 41].

This study area located in a mountainous region consisting of primary and secondary closed evergreen broadleaved lowland forests [42]. The dominant species in the study area are of families: Fagaceae, Myrtaceae, Lauraceae, Cannabaceae, Leguminosae, Dipterocarpaceae, Malvaceae, Meliaceae, Myristicaceae, Burseraceae, and Annonaceae [42, 43]. There are four main land-use type grass, natural forest, plantation forest, and agricultural land in A Luoi district. The predominant soil types are of Humic Acrisols, Hyperdystric Acrisols, Arenic Acrisols, and Ferralic Acrisols [44]. According to Thai [45], the study area consists of mainly lowland evergreen broadleaf forests. The classification system of forest-based ecosystem encloses five main stories, namely, (i) upper storey where tree with height of about 40 to 50 m; (ii) ecological dominance where most of trees belong to Fagaceae, Lauraceae, Magnoliaceae, Burseraceae, and Meliaceae; (iii) lower storey where tree height is from 8 to 15 m belonging to Annonaceae, Ulmaceaeand, Myristicaceae, and Clusiaceae; (iv) under storey where trees are from 2 to 8 m high; and (v) climbers with the height of less than two meters. Four forest types of the lowland forest consist of evergreen closed, semideciduous closed, deciduous closed and closed, and hard leaved that echoed the development of lowland tropical forests in Vietnam since 1960s. The four-forest-type classification method of Germany introduced by Loeschau to Vietnam in 1959 has been widely applied [46, 47]. The rich forest distributes in the very remote and high terrain area where forests are protected and forest structure is well-preserved, called undisturbed forest (UF), which had the basal area of about 30 m²·ha⁻¹ [48]. The UF is dominated by Fagaceae, Lauraceae, Dipterocarpaceae, Leguminosae, and Meliaceae. The second forest type is classified as a medium forest where the forest has been somehow slightly logged and disturbed, canopy fragmentation exists, and its structure is somehow maintained with the basal area ranging from 21–26 m²·ha⁻¹ [49]. This forest type is classified as less disturbed forest (LF), which is dominated by Fagaceae, Lauraceae, Euphorbiaceae, and Sapindaceae. The third type is secondary forest, where forest is heavily disturbed (DF), which is dominated with Myristicaceae, Clusiaceae, Annonaceae, Euphorbiaceae, and Myrtaceae [45, 50] with the basal area from 10–21 m²·ha⁻¹ [49].

3. Materials and Methods

3.1. Ground-Truth Samples. In order to gather the ground-truth data for (1) forest types classification and (2) assess the correlation between vegetation indices and forest vertical
and horizontal structures, we based on the local forest status map to randomly layout 90 sampled plots (30 × 33.3 m) in three different forest classes hereafter referred to as forest types of UF, LF, and DF to measure the height (H) in meter and the diameter at breast height (DBH) ≥10 in a centimeter of all living trees. Forest stand parameters were collected using conventional forest inventory techniques [51]. The tree species were first recorded in Vietnamese nomenclature and then to scientific names [52]. Names of all tree species were then checked to avoid synonym [53, 54]. The mean stand parameters were then calculated and analyzed and the t test was conducted [55, 56]. The dominance of tree species in basal area (m²/ha⁻¹), abundance of tree species occurring as number stem of a tree species (N/ha⁻¹), frequency (%), and Importance Value Index (IVI) of the most dominant and abundant species were calculated [57, 58]. The in-stored random samples in GPS were used to guide the field team to the sampled site. The minimum interval distance among plots is around 200 meters. The reference dataset has been created by a field survey where ground-truth values have

**Figure 1:** A key process on land cover classification, vegetation indices extraction, and verification of research workflow.

**Figure 2:** Research area. The ground-truth sampling plots were conducted at three forest types: UF = undisturbed forest; LF = less disturbed forest; DF = disturbed forest.
been recorded using the Global Positioning System (GPS). The ground-truth sampling was conducted from the 3rd of April to the 28th of May 2017.

3.2. Remote Sensing Data. Both Sentinel-2-Level 1C (ID: L1C_T48PYC_A009124_20170322T033236) acquired on 22 March 2017 and Landsat-8 (ID: LC08_L1TP_125049_20170809_20170823_01_T1, acquired on 09 August 2017) were downloaded from the website of the United States Geological Survey [59]. The imagery was atmospherically corrected by using the Sen2Cor tool incompatibility in the platform of the Sentinel Application Platform (SNAP) toolbox [60]. All bands, except band 10 of Sentinel-2, were converted to radiance, TOA reflectance, and surface reflectance and resampled to 10 m x 10 m resolution with the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model referencing before compositing and extracting by mask to the study area. The Sentinel-Cirrus band was neither used for land cover classification nor vegetation calculation [61]. The composited 10 m band imagery were used for classification and extraction of vegetation indices in this study. The Landsat-8 provides 11 multispectral bands, nine bands except for band 8 and band 9 (Panchromatic and Cirrus) were not used for classification. All selected bands were resampled to 10 m x 10 m resolution for sampling and classification. The flow of remote sensing data collection and process is presented in Figure 3.

3.3. Land Cover Classification Training and Testing Samples

3.3.1. Training Dataset Classification and Testing. A set of polygons was created to collect training sample data by using ArcGIS 10.6.1 toolbox. The training data were manually sampled in composited RGB imagery of Sentinel-2. To enhance the accuracy of training sample, the up-to-date imagery of Google Earth was also used as a reference. Different sizes of polygons may differ in the number of pixels per land cover class. The training and test samples are presented in Table 1.

The land use/land cover (LULC) of the selected area is divided into seven classes. In accordance with Land Law 2013 of Vietnam, the three main land-use types are agricultural, nonagricultural, and unused lands. Of the seven land cover classes in Table 1, RC and WR belong to non-agricultural land. The Agr, SB, DF, LF, and UF are of agricultural land. Forest lands are the most dominant in the current study area [62]. Land cover and land use were classified into four land-use type arable land, plantation forest, natural forest for production, and bare land [44]. Thus, the seven land cover classes are the most representative of the natural land in this area. The residential and construction areas illustrated in this study consist of civil constructions such as houses, roads, irrigation, schools, and clinics. Water area includes river which, at the imagery acquired data, was partly shallow. The agricultural area is of rice, maize, and vegetable production. Slash and burnt is either upland land which is considered as land for woodlots and exotic tree species plantation such as Acacia auriculiformis, Acacia mangium, Acacia hybrid, Hevea brasiliensis, Anacardium occidentale, and Hopea ordorata or periodically being burnt after harvesting. This type of land exists in all mountainous areas of Vietnam. The three forest cover classes classified into disturbed forests were (i) forests were heavily degraded or deforested (DF), (ii) forests were somehow degraded but its structure is maintained (LF), and (iii) forests where minor disturbances by nature or human have occurred and its structure is well preserved (UF) [47, 63].

3.3.2. Random Forest Classification. The two most important parameters in RF land cover classification are ntree and mtry which affect the results of land cover classification where the random classifier vector generates from different pixels and each tree cast is a unit vote for the most class to rank the input vector [14, 64]. Twelve bands of the Sentinel-2 as 12 input variables and the nine bands of the Landsat-8 as nine bands as input variables were used for classification. The composition of RF classification is basically from subsamples of ntree, which are trained and fit best to the most popular class. A total of 2236 pixels for seven classes as seven land-use types used unbalanced random sampling. To support the RF algorithm, a ratio of 70% training pixels (2258 training samples) was used to run the RF algorithm and 30% of the total training pixels (964 training samples) was used for testing. The remaining samples are used, by RF, to obtain the class error estimates indicating via the score of out of bag (OOB). According to some researchers, ntree parameters applied in RF classifier can be default [65, 66], and other studies state that increase of ntree will produce the better performance of accuracy estimates [65]. The number of ntree in this study was defined as a list of nine levels continuously ranging from 100, 200, 300, 400, 500, 600, 700, 800, and 900. The feature in each split called mtry which is controlled by user in the subsamples and feeds the RF classifier [6]. The default mtry for p features is mtry = √p, where p is the number of variable predictors; thus, increasing mtry may improve the performance [67]. The mtry in this study was defined (1, variables) to run the classifier with the same defined ntree (100, 200, 300, 400, 500, 600, 700, 800, and 900). Therefore, the highest accuracy of mtry and ntree result was selected. Random forest classification analysis was done in R studio.

3.3.3. Accuracy Assessment and Validation. Three forest classes are based on the field sample plots. The idea is to estimate the accuracy of quantification of mapping from using remote sensing data to the ground-truth conditions and compare the performance of different satellite images in terms of its bands as variables. We can then determine the level of error of users and producers that might be contributed by the land use or land cover in further analyses in which it is incorporated. Accuracy assessment of each defined class drives from an error matrix that compares map information with reference data and the sampled area/points; these types of errors are driven by the producer’s and user’s accuracies, respectively [68, 69]. The accuracy is
3.3.4. Extraction of Vegetation Indices. The values of ten VIs were calculated from different combinations of spectral bands, mostly from NIR and red bands, derived from the better-performed images [72–74]. Nine out of ten VIs are broadband greenness that help to understand canopy leaf area, canopy formation, and vegetation productivity.

These VIs provide and compare the reflectance peak in the range of near-infrared to red band. While selection of Atmospherically Resistant Vegetation Index (ARVI) was as a test to measure the reflectance of the narrow and steep slope of vegetation structure change compared with broadband greenness with the hope that it can help to differentiate different vegetation canopy structures.

The VIs were extracted from both ground-truth plots and training sample points. The Normalized Difference Vegetation Index (NDVI) is a simple and effective index that helps to quantify the green vegetation, ranging from −1 to 1, and the indicated green healthy and dense vegetation is between 0.2 and 0.8 [75, 76]. The Infrared Percentage Vegetation Index (IPVI) is a nonnegative vegetation index.
that helps to measure the percentage of the radiance of both
near-infrared and red bands [77, 78].

Similar to NDVI, the Green Normalized Difference
Vegetation Index (GNDVI) is more sensitive to chlorophyll
concentration than NDVI and used for measuring the green
spectrum from 540 to 570 nm instead of the red spectrum
[79]. The Atmospherically Resistant Vegetation Index
(ARVI) resists to atmospheric effects compared with NDVI
and is supported by a self-corrected process on the red band
[80]. The range of ARVI is dynamically similar to that of
NDVI, but it is four times less sensitive to atmospheric as
NDVI. The Enhanced Vegetation Index (EVI) values range
from -1 to 1, of which the healthy vegetation denotes from
0.2 to 0.8 [31, 81]. The Normalized Difference Index (DVI) is
sensitive to vegetation and more distinguishing with soil and
is more linear and of use for vegetation cover monitoring
[82, 83]. The Normalized Difference Index (NDI45) with less
saturation at higher values than the NDVI is the ratio be-
tween the Red Edge (band 5) and the red band (Band 4) [84].
The value of the Ratio Vegetation Index (RVI) ranges from 0
to 30 in which values between 2 and 8 indicate health
vegetation. The RVI reduces the effects of atmosphere and
topography [85–87]. This effect occurs in the natural surface
in the mountainous area. The Perpendicular Vegetation
Index (PVI) is a generation of DVI differentiating soil lines of
different slopes and near-infrared (NIR) axis in degree and
is to some extent sensitive to atmospheric variations
[88]. The algorithm of Transformed Normalized Difference
Vegetation Index (TNDVI) is the square root of NDVI and
indicates a relationship between green biomass that is found
in a pixel [89].

In this study, we calculated ten VIs from the best-per-
formed image in Table 2 from 90 ground-truth randomly
sampled plots (GTVs) and 1303 training sample points of
three substratified forest types (CLVs) in Table 3. T-tests of
the mean were done among three forest types of both GTVs
and LCVs and correlation was checked among VIs of the
GTVs and horizontal and vertical forest structure [90] to
test the significant difference among forest types and
ground-truth and training sample plots. All T-tests of forest
parameters and vegetation indices were tested by using
Statistica 13.5.

4. Results

4.1. Ground-Truth Input. The forest sampled parameters of
three forest types UF, LF, and DF are not uniquely different.
The mean of height as a vertical structure of forest stands LF
and DF is significant different from UF at \( p < 0.05 \) but not
between LF and DF. The mean basal area and of volume as a
horizontal structure of forest stand is significantly different
among UF, LF, and DF while tree stems per plot were not
significantly different among forest types in Table 4.

The composition of species and families in different
forest types illustrated the mean height, dominance (m\(^2\)/ha),
abundance (N/ha), frequency (%), and Importance Value
Index (IVI) of each forest type in Tables 5 and 6. Tables 5
shows the most dominant, abundant species in each forest
type with its frequency and IVI in which \( L. \ ducampii,\)

A. grandiflora, T. myriocarpa, and S. roxburghii in UF where
\( L. \ ducampii \) 59 stems were found over 20 plots occupying
19.67 trees per ha while 39 stems of \( S. \ roxburghii \) were found
in 21 plots and were the most dominant with 2.31 m\(^3\)/ha.
This species is also dominant in the height of 21.98 m with
70% frequency. In LF, \( B. \ tonkinensis \) was found most
dominant, but \( L. \ ducampii \) was also found most abundant.
However, \( S. \ macropodum \) and \( M. \ mediocris \) are at the lower
height stratum compared with \( L. \ ducampii \) and \( B. \ tonkinensis. \ \( L. \ ducampii \) is more abundant but less
dominant than \( B. \ tonkinensis. \ In DF, \( G. \ oliveri, \ K. \ furfuracea, \ E. \ petelotii, \ and \ S. \ lanceolatum \ are the four
most dominant and abundant species. Of those four species,
\( G. \ oliveri \) and \( K. \ furfuracea \) are the two most dominant and
abundant species.

Table 6 describes the species which prevails in the higher
canopy stratum of which \( D. \ kerrii \) has the mean height of
27.98 m followed by \( T. \ bellirica \) with 26.70 m, while
\( S. \ roxburghii \) and \( D. \ grandiflorus \) are most dominant and
abundant in UF, where the three out of four dominant and
abundant species are of Dipterocarpaceae and one species is
of Combretaceae. Furthermore, \( S. \ roxburghii \) and
\( D. \ grandiflorus \) were found most frequent with 39 stems over
21 sampled plots and 17 out of 30 sampled plots with 34
stems, respectively. The number of species in LF and the
dominance and abundance of four species \( F. \ lacor, \ S. \ roxburghii, \ C. \ indica, \ and \ T. \ myriocarpa \ are not so much
varied. \( F. \ lacor, \) the highest vertical stratum, has four stems
distributing in four out of 30 sampled plots followed
\( S. \ roxburghii \) which is 21.27 m in height with eight stems
distributed in six sampled plots. The other two species
\( C. \ indica \) and \( T. \ myriocarpa \) are more abundant than two
other species mentioned above. The vertical stratum in LF is
even more scattered than the species in UF and LF. \( P. \ ellipticum, \) one single stem distributed on one sampled plot,
was found to be the highest in this forest type with 34.50 m.
The height of three species \( A. \ chinense, L. \ verticillata, \) and
\( A. \ pilosa \) ranges from 18.06 m to 19.17 m.

The four most horizontal dominant and abundant
families are Fagaceae, Lauraceae, Leguminosae, and Di-
pterocarpaceae. Of those, Fagaceae and Lauraceae are more
abundant with 52.7 stems and 33.3 stems per hectare while
Dipterocarpaceae is the most dominant with 93 stems
recorded in 29 out of 30 sampled plots. Fagaceae is more
dominant and abundant than in LF, but a shift is with
Lauraceae which is more dominant and abundant in LF.
Euphorbiaceae and Myrtaceae were not found dominant
and abundant in UF but in LF and DF. It is obvious that the
Myristicaceae and Clusiaceae are only found dominant and
abundant in DF. The vertical tree families in UF are
Leguminosae and Dipterocarpaceae of which Dipter-
ocarpaceae is preheight dominant and abundant. The
dominance and abundance of Combretaceae are more than
Moraceae. Compared with the horizontal family structure,
the dominant and abundant families belong to Dipter-
ocarpaceae, Combretaceae, Moraceae, Leguminosae, Bur-
eraceae, and Polygalaceae. Of those, Dipterocarpaceae and
Leguminosae were found in both vertical and horizontal
family structures of species composition in the study area.
Table 2: Vegetation radiometric indices.

<table>
<thead>
<tr>
<th>Vegetation indices</th>
<th>Formula</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI</td>
<td>((B_8 - B_4)/(B_8 + B_4))</td>
<td>[82]</td>
</tr>
<tr>
<td>IPVI</td>
<td>(IPVI = (B_8 - (B_8 + B_4)))</td>
<td>[79]</td>
</tr>
<tr>
<td>GNDVI</td>
<td>((B_8 - B_3)/(B_8 + B_3))</td>
<td>[92]</td>
</tr>
<tr>
<td>ARVI</td>
<td>(B_8 - [B_4 - \gamma(B_4 - B_2)/B_8 + [B_4 - \gamma(B_4 - B_2)])</td>
<td>[80]</td>
</tr>
<tr>
<td>EVI</td>
<td>((G \ast ((B_8 - B_4)/(B_8 + C_1 \ast B_4 - C_2 \ast B_2 + L))) (\gamma = \text{gamma} = 1)</td>
<td>[93]</td>
</tr>
<tr>
<td>ND145</td>
<td>((B_5 - B_4)/(B_5 + B_4))</td>
<td>[84]</td>
</tr>
<tr>
<td>RVI</td>
<td>(B_8/B_4)</td>
<td>[85]</td>
</tr>
<tr>
<td>DVI</td>
<td>(B_8 - B_4)</td>
<td>[82]</td>
</tr>
<tr>
<td>PVI</td>
<td>((ap \ast B_8 - bp \ast B_4))</td>
<td>[88]</td>
</tr>
<tr>
<td>TNDVI</td>
<td>(\sqrt{(B_3 - B_4)/(B_3 + B_4)} + 0.5)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Mean vegetation indices of forest types in ground-truth and training sample plots extracted from Sentinel-2 imagery dated on 22 March 2017.

<table>
<thead>
<tr>
<th>Vegetation indices</th>
<th>Ground-truth plots (GTVIs)</th>
<th>Training sample plots (LCVIs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(UF (N = 30))</td>
<td>(LF (N = 30))</td>
</tr>
<tr>
<td>NDVI</td>
<td>0.84 ± 0.02a(^A)</td>
<td>0.85 ± 0.01a(^A)</td>
</tr>
<tr>
<td>IPVI</td>
<td>0.92 ± 0.01a(^A)</td>
<td>0.93 ± 0.01a(^A)</td>
</tr>
<tr>
<td>GNDVI</td>
<td>0.73 ± 0.03a(^A)</td>
<td>0.74 ± 0.02a(^A)</td>
</tr>
<tr>
<td>ARVI</td>
<td>0.85 ± 0.03a(^A)</td>
<td>0.84 ± 0.02a(^A)</td>
</tr>
<tr>
<td>EVI</td>
<td>0.22 ± 0.01a(^A)</td>
<td>0.33 ± 0.05a(^A)</td>
</tr>
<tr>
<td>ND145</td>
<td>0.54 ± 0.07a(^A)</td>
<td>0.53 ± 0.04a(^A)</td>
</tr>
<tr>
<td>RVI</td>
<td>12.1 ± 1.9a(^A)</td>
<td>12.8 ± 1.2a(^B)</td>
</tr>
<tr>
<td>DVI</td>
<td>0.07 ± 0.02a(^A)</td>
<td>0.11 ± 0.02a(^A)</td>
</tr>
<tr>
<td>PVI</td>
<td>0.12 ± 0.03a(^A)</td>
<td>0.20 ± 0.03a(^A)</td>
</tr>
<tr>
<td>TNDVI</td>
<td>1.07 ± 0.02a(^A)</td>
<td>1.10 ± 0.01a(^A)</td>
</tr>
</tbody>
</table>

Within rows, values followed by the same lowercase letter (a, b) are not significantly different \((p < 0.05)\) between forest types; Within rows, values followed by the same capital letter (A, B) are not significantly different \((p < 0.05)\) of forest types between ground-truth and training sample plots.

Table 4: Parameters of different forest types.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>UF ((N = 30))</th>
<th>LF ((N = 30))</th>
<th>DF ((N = 30))</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of plots ((1000 m^2))</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Mean height ((H)) in meter</td>
<td>16.4 ± 1.2a</td>
<td>14.7 ± 1.2b</td>
<td>14.3 ± 0.9bc</td>
</tr>
<tr>
<td>Basal area ((\text{BA})) ha(^{-1})</td>
<td>33.1 ± 0.6a</td>
<td>21.3 ± 0.3b</td>
<td>15.6 ± 0.4c</td>
</tr>
<tr>
<td>Volume ((V)) ha(^{-1})</td>
<td>361 ± 8.8a</td>
<td>211 ± 4.6b</td>
<td>147 ± 4.7c</td>
</tr>
<tr>
<td>Tree stem ha(^{-1})</td>
<td>445 ± 9.4a</td>
<td>466 ± 13.5a</td>
<td>400 ± 12.0a</td>
</tr>
</tbody>
</table>

The same letter a, b, or c indicates no significant difference at \(p < 0.05\).

4.2. Performance of RF Classifier. The best-produced results from mtry and ntree combined-datasets showed the highest accuracy presented in Figure 4. Tuning results of two RF classifiers illustrated repeated cross-validation accuracy where mtry = 2 at ntree = 500 in Landsat and mtry = 5 at ntree = 700 in Sentinel-2.

The same mtry and training sampled dataset which were used to assess the error of each land-use type in Figure 5. The results showed the OOB in Landsat-8 of 31.6% in Landsat, and the OOB is of 14.3% in Sentinel-2.

4.3. Comparison of Sensors over Class Validation and Assessment. The agreement of the RF classification including the overall accuracy and the kappa statistic is presented in Tables 7 and 8. The accuracy assessment of land cover RF classifiers in Landsat-8 showed the overall accuracy of 68% and Kappa of 67.5%. The results showed that the producer accuracy of Water and River (WR) and Slash and Burnt (SB), which are presented in Table 7, has the lowest accuracy with 20% and 1%, respectively.

The overall accuracy of the RF classifier in Sentinel-2 in Table 8 was 86% and kappa of 83%. The producer’s error and user’s error of water and river class in Sentinel-2 are the highest compared with other classes.

Figure 6 showed pairwise matrix of 7 land cover classes in which the most misinterpreted training samples are of
Table 5: Dominance and abundance species of horizontal structure in different forest types.

<table>
<thead>
<tr>
<th>Types</th>
<th>Species</th>
<th>Mean $H$ (m)</th>
<th>Dominance ($m^2/ha$)</th>
<th>Abundance (N/ha)</th>
<th>Frequency (%)</th>
<th>IVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>UF</td>
<td>Lithocarpus ducampi (Hickel and A. Camus) (59 : 20)</td>
<td>16.97</td>
<td>1.51</td>
<td>19.67</td>
<td>66.67</td>
<td>7.58</td>
</tr>
<tr>
<td></td>
<td>Aphananthis grandiflora (50 : 19)</td>
<td>15.67</td>
<td>0.98</td>
<td>16.67</td>
<td>63.33</td>
<td>6.70</td>
</tr>
<tr>
<td></td>
<td>Terminalia myriocarpa Van Heurck and Müll. Arg (41 : 19)</td>
<td>20.23</td>
<td>1.58</td>
<td>13.67</td>
<td>63.33</td>
<td>7.49</td>
</tr>
<tr>
<td></td>
<td>Shorea roxburghii G. Don (39 : 21)</td>
<td>21.98</td>
<td>2.31</td>
<td>13.00</td>
<td>70.00</td>
<td>8.74</td>
</tr>
<tr>
<td></td>
<td>Others (1148 : 96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>26.69</td>
<td>382.67</td>
<td></td>
<td></td>
<td>269.49</td>
</tr>
<tr>
<td>LF</td>
<td>Lithocarpus ducampi (Hickel and A. Camus) (54 : 16)</td>
<td>16.44</td>
<td>1.06</td>
<td>18.00</td>
<td>53.33</td>
<td>6.51</td>
</tr>
<tr>
<td></td>
<td>Schiaphtum macrodorum (Maq) Beemsee ex K. Heyne (41 : 21)</td>
<td>14.03</td>
<td>0.39</td>
<td>13.67</td>
<td>70.00</td>
<td>6.07</td>
</tr>
<tr>
<td></td>
<td>Magnolia medocis (Dandy) Figlar (41 : 18)</td>
<td>14.14</td>
<td>0.63</td>
<td>13.67</td>
<td>60.00</td>
<td>5.91</td>
</tr>
<tr>
<td></td>
<td>Bersera tonkinensis Guillaum (40 : 19)</td>
<td>18.53</td>
<td>1.61</td>
<td>13.33</td>
<td>63.33</td>
<td>7.95</td>
</tr>
<tr>
<td></td>
<td>Others (1221 : 107)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.70</td>
<td>407.00</td>
<td></td>
<td></td>
<td>273.56</td>
</tr>
<tr>
<td>DF</td>
<td>Garcinia oliveri Pierre (66 : 25)</td>
<td>12.49</td>
<td>0.49</td>
<td>22.00</td>
<td>83.33</td>
<td>8.01</td>
</tr>
<tr>
<td></td>
<td>Scaphium macropodium (Miq) Beumee ex K. Heyne (48 : 21)</td>
<td>14.76</td>
<td>0.54</td>
<td>16.00</td>
<td>70.00</td>
<td>7.12</td>
</tr>
<tr>
<td></td>
<td>Enicosanthellum petelotii (Merr.) Ban (46 : 17)</td>
<td>12.25</td>
<td>0.30</td>
<td>15.33</td>
<td>56.67</td>
<td>5.64</td>
</tr>
<tr>
<td></td>
<td>Syzygium lanceolatum (Lam.) Wight and Arn (45 : 22)</td>
<td>13.00</td>
<td>0.42</td>
<td>15.00</td>
<td>73.33</td>
<td>7.12</td>
</tr>
<tr>
<td></td>
<td>Others (997 : 105)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.84</td>
<td>332.33</td>
<td></td>
<td></td>
<td>272.12</td>
</tr>
</tbody>
</table>

Species presented in bracket, i.e, 59 : 20 indicate the number of individual stems (N) representing the number of corresponding plots where the species are dominated and abundant with its occurrence frequency; IVI = Importance Value Index of the species in the study area.

Table 6: Dominance and abundance species of vertical structure in different forest types.

<table>
<thead>
<tr>
<th>Types</th>
<th>Species</th>
<th>Mean $H$ (m)</th>
<th>Dominance ($m^2/ha$)</th>
<th>Abundance (N/ha)</th>
<th>Frequency (%)</th>
<th>IVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>UF</td>
<td>Diptocerox kerrii King (6 : 3)</td>
<td>27.98</td>
<td>0.51</td>
<td>2.00</td>
<td>10.00</td>
<td>5.19</td>
</tr>
<tr>
<td></td>
<td>Terminalia bellirica (Gaertn.) Roxb. (4 : 4)</td>
<td>26.70</td>
<td>0.29</td>
<td>1.33</td>
<td>13.33</td>
<td>2.97</td>
</tr>
<tr>
<td></td>
<td>Shorea roxburghii G. Don (39 : 21)</td>
<td>21.98</td>
<td>2.31</td>
<td>13.00</td>
<td>70.00</td>
<td>8.74</td>
</tr>
<tr>
<td></td>
<td>Diptocerox grandiflorus Blanco (34 : 17)</td>
<td>21.04</td>
<td>1.81</td>
<td>11.33</td>
<td>56.67</td>
<td>7.51</td>
</tr>
<tr>
<td></td>
<td>Others (1254 : 96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>28.15</td>
<td>418.00</td>
<td></td>
<td></td>
<td>275.60</td>
</tr>
<tr>
<td>LF</td>
<td>Ficus lacor Buch. Ham. (4 : 4)</td>
<td>24.50</td>
<td>0.33</td>
<td>1.33</td>
<td>13.33</td>
<td>3.97</td>
</tr>
<tr>
<td></td>
<td>Shorea roxburghii G. Don (8 : 6)</td>
<td>21.27</td>
<td>0.47</td>
<td>2.67</td>
<td>20.00</td>
<td>4.39</td>
</tr>
<tr>
<td></td>
<td>Castanopsis indica (Roxb. ex Lindl.) A.DC (11 : 6)</td>
<td>20.12</td>
<td>0.34</td>
<td>3.67</td>
<td>20.00</td>
<td>3.59</td>
</tr>
<tr>
<td></td>
<td>Terminalia myriocarpa Van Heurck and Müll. Arg (10 : 5)</td>
<td>19.42</td>
<td>0.26</td>
<td>3.33</td>
<td>16.67</td>
<td>3.19</td>
</tr>
<tr>
<td></td>
<td>Others (1364 : 107)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.98</td>
<td>454.67</td>
<td></td>
<td></td>
<td>284.87</td>
</tr>
<tr>
<td>DF</td>
<td>Placolobium ellipticum N. D. Khoi and Yakovlev (1 : 1)</td>
<td>34.50</td>
<td>0.12</td>
<td>0.33</td>
<td>3.33</td>
<td>5.40</td>
</tr>
<tr>
<td></td>
<td>Alangium chinense (Lour.) Harms (3 : 2)</td>
<td>19.17</td>
<td>0.10</td>
<td>1.00</td>
<td>6.67</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>Litsea verticillata Hance (6 : 3)</td>
<td>18.07</td>
<td>0.11</td>
<td>2.00</td>
<td>10.00</td>
<td>2.42</td>
</tr>
<tr>
<td></td>
<td>Actinodaphne pilosa (7 : 5)</td>
<td>18.06</td>
<td>0.23</td>
<td>2.33</td>
<td>16.67</td>
<td>3.40</td>
</tr>
<tr>
<td></td>
<td>Others (1185 : 105)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15.02</td>
<td>395.00</td>
<td></td>
<td></td>
<td>286.10</td>
</tr>
</tbody>
</table>

Species presented in bracket, i.e, 6 : 3, indicate the number of individual stems (N) representing the number of corresponding plots where the species are dominated and abundant with its occurrence frequency; IVI = Importance Value Index of the species in the study area.

water and river (WR) and slash and burst (SB) classes in Landsat-8 and WR in Sentinel-2. Interpretation of training samples in Sentinel-2 is more constant. In Figure 7, the water body and slash and burnt ratios are more visualized in Landsat-8 than those in Sentinel-2.

4.4. Difference of Vegetation Indices. Ten extracted vegetation indices from three forest cover classes of the best-performed image are presented in Table 2 which consists of VIs values of GTVIs and CLVIs. T tests of the mean were done among three forest types of both GTVIs, and LCVIs in Table 3. The NDVI, IPVI, EVI, DVI, PVI, and TNDVI in different forest types of both GTVIs and LCVIs showed significant difference at $p < 0.05$. However, GNDVI, ARVI, ND145, and RVI of different forest types of GTVIs were not significantly different. Not all vegetation indices of each corresponding forest type in GTVIs and LCVIs are significant different. The vegetation indices of EVI, DVI, and TNDVI of GTVIs and LCVIs in each corresponding forest type are significantly different at $p < 0.05$.

4.5. Correlation of Vegetation Indices. The correlation and regression among VIs of the GTVIs and horizontal and vertical forest structure are presented in Table 9 and Figure 8. The figures presented in Table 9 do not only describe the correlation between the height and BA with ten vegetation indices but also among themselves for further understandings. In this study, we decided to choose the correlations more than 50% with BA and above 40% with height and thus focused on the correlation between BA and height with EVI, PVI, DVI, and TNDVI. The correlation of these four VIs with BA ranges from 0.63 to 0.66, while the
Table 7: Accuracy assessment of land cover classes of Landsat-8.

<table>
<thead>
<tr>
<th>LC/LU</th>
<th>RC</th>
<th>WR</th>
<th>Agr</th>
<th>SB</th>
<th>DF</th>
<th>LF</th>
<th>UF</th>
<th>User accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>104</td>
<td>19</td>
<td>13</td>
<td>30</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>WR</td>
<td>8</td>
<td>17</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>53</td>
</tr>
<tr>
<td>Agr</td>
<td>30</td>
<td>37</td>
<td>112</td>
<td>48</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td>SB</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>DF</td>
<td>0</td>
<td>2</td>
<td>13</td>
<td>9</td>
<td>134</td>
<td>5</td>
<td>0</td>
<td>82</td>
</tr>
<tr>
<td>LF</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>118</td>
<td>16</td>
<td>75</td>
</tr>
<tr>
<td>UF</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>14</td>
<td>170</td>
<td>86</td>
</tr>
<tr>
<td>Producer accuracy</td>
<td>72</td>
<td>20</td>
<td>75</td>
<td>01</td>
<td>86</td>
<td>81</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Accuracy assessment of land cover classes of Sentinel-2.

<table>
<thead>
<tr>
<th>LC/LU</th>
<th>RC</th>
<th>WR</th>
<th>Agr</th>
<th>SB</th>
<th>DF</th>
<th>LF</th>
<th>UF</th>
<th>User accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>108</td>
<td>13</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>82</td>
</tr>
<tr>
<td>WR</td>
<td>2</td>
<td>52</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>Agr</td>
<td>11</td>
<td>11</td>
<td>123</td>
<td>8</td>
<td>6</td>
<td>0</td>
<td>4</td>
<td>75</td>
</tr>
<tr>
<td>SB</td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>79</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>79</td>
</tr>
<tr>
<td>DF</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>180</td>
<td>9</td>
<td>0</td>
<td>93</td>
</tr>
<tr>
<td>LF</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>128</td>
<td>8</td>
<td>93</td>
</tr>
<tr>
<td>UF</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>158</td>
<td>88</td>
</tr>
<tr>
<td>Producer accuracy</td>
<td>85</td>
<td>58</td>
<td>84</td>
<td>84</td>
<td>97</td>
<td>84</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kappa</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 6: Explanation of each class prediction of (a) Landsat-8 and (b) Sentinel-2 using square matrix. The higher the result of reference and prediction, the better the accuracy of classification (classes are 1 = RC; 2 = WR; 3 = Agr; 4 = SB; 5 = DF; 6 = LF; and 7 = UF).

Figure 7: Classification map by random forest presenting results of the best accuracy of classifiers in (a) Landsat-8 and (b) Sentinel-2, respectively. Forest area is most dominant in the study area with 88% of the total natural land [62].
correlation of height with these four VIs ranges from 0.43 to 0.49. Three VIs showed the lowest correlation with BA, and heights were GNDVI, ARVI, and NDI45.

4.6. Relationship of VIs with Horizontal and Vertical Structures of Dominance Species. Species that are mostly horizontally dominant and abundant showed a correlation with VIs in UF of Terminalia myriocarpa with 41 stems distributed in 19 sampled plots as presented in Table 5 followed by Lithocarpus ducampii. This species is the most abundant in UF and LF, but its horizontal parameter was not as correlational as this in UF, but Bursera tonkinensis showed most correlation with VIs followed by Magnolia mediocris. In DF, Enicosanthellum petelotii showed more correlation with VIs than Garcinia oliveri, Syzygium lanceolatum, and Knema furfuracea in Table 10.

In the vertical structure of most dominant and abundant tree species presented in Table 6, a strong correlation with VIs was shown. Of which, in UF, Dipterocarpus kerrii is followed by Terminalia bellirica, Dipterocarpus grandiflorus, and Shorea roxburghii. In LF, Shorea roxburghii showed a relatively higher correlation with VIs than Ficus laco, Castanopsis indica, and Terminalia myriocarpa. While in DF, those four species Placolobium ellipticum, Alangium chimense, Litsea verticillata, and Actinodaphne pilosa showed a strong correlation with VIs presented in Table 11.

4.7. VIs Regression with Horizontal and Vertical Structures. Negative linear regression of four vegetation indices with the mean basal area and height of the ground-truth sample plots in the study area is presented in Figure 8. As presented above, we focused on the regression of BA and height with EVI, PVI, DVI, and TNDVI. As a result, the regression of these four VIs is better with the horizontal structure of forest stand than that with vertical ones. The results showed that EVI had the negative linear regression with the basal area than DVI, PVI, and TNDVI with \( R^2 \) of 0.44, 0.43, 0.43, and 0.40, respectively. On the other hand, the TNDVI had the negative linear regression with the mean height of forest stand than EVI, PVI, and DVI with \( R^2 \) values of 0.24, 0.21, 0.19, and 0.18, respectively.

5. Discussions

5.1. Species’ Vertical and Horizontal Structures of Different Forest Types. Referencing the local forest status maps before setting up random sample plots could hint for a significant difference in the horizontal structure of different forest types. The difference shows the reliability of the preexisting forest status map. The vertical stratum of forest types in Table 4 is not unique due to either human disturbances or geographical factors such as elevations, slopes, and species distribution [43, 54, 94]. The forest type (UF) is composed of more D. kerrii, T. bellirica, S. roxburghii, and D. grandiflorus species that belong to Dipterocarpaceae, Leguminosea, Combretaceae, Moraceae, and Burseraceae families. These families usually distribute in the evergreen or semievergreen forests [43, 50, 95, 96]. These species are more vertical dominant than those in LF that consists of either the same species or families. Species and families, classified as a high vertical stratum, in DF are less dominant and abundant than those in UF and LF. The mix of forest species distribution and forest destruction in LF and DF causes no significant difference in the mean of vertical stratum [46, 97, 98] in Table 6. The mixture of forest type-based dominant species between UF and LF is of evergreen and semievergreen forest types; these dominant and abundant species and families distribute at the elevation of less than 900 m above sea level where the mean annual rainfall is more than 2500 mm [45, 99]. In contrast, the similarity of the number of stems, species, and families in three different forest types could be due to the intermixed location of sample plots and disturbances in forest type, respectively, where the mean of those three parameters shows no significant difference [100–102]. On the other hand, the species and families contributing to the horizontal structure of forest stands in the tropic of Malaysia-Indonesia and India regions are more light-demanding species of L. ducampii, A. grandiflora, T. myriocarpa, S. roxburghii, S. macropodum, M. mediocris, B. tonkinensis, G. oliveri, K. furfuracea, E. petelotii, and S. lanceolatum belonging to Fagaceae, Lauraceae, Leguminosae, Dipterocarpaceae, Euphorbiaceae, Myrtaceae, and Clusiae [42, 45, 54, 63, 103–105].

5.2. Performance of RF Classifier. There have been contradictions on the performance of RF classifier on various training sample sizes with different satellite images in many studies [64]. The proportion of the land cover types of different land uses in the mountainous region is imbalanced and sometimes fragmented into small scale areas that led to a nonunique number of training pixels in classes in Table 1. The smaller number of pixels in each training dataset of each class outputted the smaller ratio of producer accuracy resulting in the precision of user’s accuracy, but the best random split and selection of subdataset of attribute at each node from sampled points by RF introduced by [14] is a successful ensemble approach. Also, it was due to the mixed cultivation of crops in the agricultural (Agr) area as well as the cultivation period in slash and burn (SB); the accuracy of producer and user was lower than that of other forest cover classes in Tables 7 and 8. Furthermore, the misinterpreted training dataset of WR and SB class in Landsat-8 is much higher than that of other classes and in Sentinel-2 in Figure 6. On the other hand, when the training sample size was unique and big enough, the classifier is less sensitive. As mentioned above, different cultivation periods plus different times of image capture result in the performance of classification [106]. The results of forest type (DF, LF, and UF) classification by the RF classifier presented in Table 8 and illustrated in Figure 6 of Sentinel-2 showed the producer’s accuracy between 84% and 97%, and the user’s accuracy was from 88% to 93%. The user’s class error was from 7% to 13%, and the producer’s error ranged from 3% to 16%. It is clear
Table 9: Correlation of vegetation indices with BA and height in ground-truth plots.

<table>
<thead>
<tr>
<th>BA (m²)</th>
<th>H (m)</th>
<th>NDVI</th>
<th>IPVI</th>
<th>GNDVI</th>
<th>ARVI</th>
<th>EVI</th>
<th>NDI45</th>
<th>RVI</th>
<th>PVI</th>
<th>DVI</th>
<th>TNDVI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.73*</td>
<td>−0.34*</td>
<td>−0.32*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.34*</td>
<td>−0.32*</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.08</td>
<td>−0.19</td>
<td>0.69*</td>
<td>0.69*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.02</td>
<td>−0.12</td>
<td>0.81*</td>
<td>0.81*</td>
<td>0.49*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.66*</td>
<td>−0.46*</td>
<td>0.60*</td>
<td>0.60*</td>
<td>0.30*</td>
<td>0.19</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.12</td>
<td>−0.13</td>
<td>0.52*</td>
<td>0.52*</td>
<td>0.27*</td>
<td>0.49*</td>
<td>0.13</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.36*</td>
<td>−0.30*</td>
<td>0.97*</td>
<td>0.97*</td>
<td>0.63*</td>
<td>0.80*</td>
<td>0.61*</td>
<td>0.54*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.65*</td>
<td>−0.43*</td>
<td>0.59*</td>
<td>0.59*</td>
<td>0.29*</td>
<td>0.19</td>
<td>0.99*</td>
<td>0.14</td>
<td>0.61*</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.65*</td>
<td>−0.43*</td>
<td>0.59*</td>
<td>0.59*</td>
<td>0.29*</td>
<td>0.19</td>
<td>0.99*</td>
<td>0.14</td>
<td>0.61*</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>−0.63*</td>
<td>−0.49*</td>
<td>0.73*</td>
<td>0.73*</td>
<td>0.44*</td>
<td>0.32*</td>
<td>0.96*</td>
<td>0.17</td>
<td>0.71*</td>
<td>0.94*</td>
<td>0.94*</td>
<td>1.00</td>
</tr>
</tbody>
</table>

BA = basal area = ((DBH²·π)/4) is the cross-sectional area in m² of a tree measured at breast height and expressed as per unit of land area. BA in Table 4 is the mean cross-sectional area in ground-truth sample plot; *Correlations are significant at $p < 0.05$. 

Figure 8: Continued.

(a)

(b)
that the accuracy does not only depend on the number of training samples but also depend on the size of the class [64, 107, 108].

5.3. Comparison of Ground-Truth Forest Cover with Those of Training Sampled Points. The detection accuracy of vegetation cover using satellite images depends on the
quality of ground-truth samplings; ten vegetation indices used in this study are to determine the difference between ground-truth plots and the training sample plots. Gebrezgabher et al. [108, 109] used multispectral of Landsat-1, Landsat-3, Landsat-5, and Landsat-7 ETM to extract NDVI, VCP, SAVI, CVA, and MNDWI. Gerstmann et al. [110] used multispectral RapidEye imagery to extract seven vegetation indices. Kobayashi et al. [111] used spectral indices of multibands of the Sentinel-2A image to improve the classification accuracy of crops. The significant difference of VIs in different forest types and each type of both GTVIs and LCVIs is presented in Table 3. (The values of seven VIs, in different forest types and each type of both GTVIs and LCVIs is presented in Table 3. The values of seven VIs, namely, NDVI, IPVI, EVI, RVI, DVI, PVI, and TNDVI, derived from ground-truthing sample plots and from training sampled points were significantly different at p < 0.05. It meant that these vegetation indices derived from UF, LF, and DF were significantly different at p < 0.05. This significant difference firstly supports to confirm that the vegetation indices derived from UF, LF, and DF were significantly different at p < 0.05. This significant difference firstly supports to confirm that the corresponding with training sample points and its classified accuracies. Only three VIs, namely, GRVI, ARVI, and DNI45, did not show any difference among different defined and classified forest types. In contrast, most of the t test results of the same value of each vegetation index (VI) from ground-truth and training sample plots of three forest types showed significant differences. This can be explained by the imbalanced number of samples in ground-truth and training sample data causing the different mean values of VI, respectively.

5.4. Relationship of VIs with Vertical and Horizontal Forest Structures. Seven out of ten VIs showed negative-significant differences with basal area of forest sample plots as horizontal structure and height as vertical structure of the forest sample plots in Table 11. The study of dos Reis et al. [112] used Landsat-STM to estimate basal area and volume of Eucalyptus camaldulensis Dehn plantation forest; the correlation of basal area and volume of the E. camaldulensis is 0.91 and −0.52, respectively. Chrysafis et al. [113] assessed the growing stock volume in Rhodopes of Greece by using Landsat-8 (OLI and Sentinel-2 to extract NDVI, DVI, CTVI, PVI, EVI, and TSAVI). Their results of the correlation of DVI, EVI, PVI, and NDVI with growing stock volume were r = −0.55, r = −0.56, r = −0.56, and r = −0.36, respectively. The study of Gamon et al. [114] showed both negative and positive correlations between NDVI and moisture vegetation indices with parameters of forests in dry and humid seasons. Vegetation indices showed strong correlations with the content of leaf chlorophyll; the correlation ratio depends on species with similar color [115]. The dominant and abundant species in Tables 10 and 11 show different correlations with BA and height of species and even the same species in different forest strata since the horizontal structure, species composition, diameter, and height class distribution in unique forest types are different [116]. For the whole forest stand where the basal area increased, the vegetation decreased. This trend defined the relationship between biomass productivity of mature forest and chlorophyll contents in the immature forest types.

6. Conclusions

The multiple spectral bands for land cover and land use classification by using random forest algorithm in Sentinel-2 showed higher accuracy than that of Landsat-8. Sentinel-2 images showed high potential for landscapes and forest type classification for conservation and management purposes in tropical lowland forests. A time series classification by using Sentinel-2 with ground-truth samples is seen as a bright hint to distinguish natural forest areas where the Sentinel-2 is available. This contributes to forest land cover mapping since the random forest classifier showed more consistency among land cover classes in Sentinel-2.

Seven vegetation indices extracted from Sentinel-2 showed significant differences between different classified forest types of ground-truth and training sample plots. The four defined vegetation indices, namely, EVI, DVI, PVI, and
TNDVI, are derived from the reflectance of the forest canopy by red and near-infrared bands which are considered as useful and possible indicators to assess the canopy horizontal structure of lowland forest, but further studies are needed for vertical structure assessment of lowland forest canopy.

The correlation between the chlorophyll content of different dominant and abundant species in different forest types suggests further studies to trait the tree species distribution and composition and to assess growth productivities of natural forest landscapes on the targeted forest tree species. The EVI, PVI, DVI, and TNDVI extracted from Sentinel-2 could help predict the better correlation with the horizontal structure but not with the vertical forest structure of forest stands.

Further research consists of more ground-truth sample plots that need to affirm the correlation and regressions between vegetation indices-based forest type with its horizontal and vertical structures.

Data Availability

The ground-truth data and satellite images are available upon request from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

We would like to express our thanks to the German Academic Exchange Service (DAAD) and the German Research Foundation and the Open Access Fund of the Göttingen University for funding this research and publication; the Department of Climate Change under the Ministry of Natural Resource and Environment of Vietnam and the University of Hue Agriculture and Forestry for their help in arranging administrative permission to access the research areas; and the experts of Sub-Forest Inventory and Forest Planning in Thua Thien Hue province; we also thank the forest experts of the Forest Management Board of the A Luoi and Rangers of Sao La Nature Reserve for their assistance during different phases of fieldwork; the botanists for tree species identification; and local authorities and local people in the A Luoi district. We thank the Department of Cartography, GIS, and Remote Sensing, Göttingen University, for providing the ArcGIS software in the framework of research.

References


[51] S. A. Sader, R. B. Waide, W. T. Lawrence, and A. T. Joyce, “Tropical forest biomass and successional age class relationships to a vegetation index derived from Landsat TM


