Large-Scale Mapping of Tree-Community Composition as a Surrogate of Forest Degradation in Bornean Tropical Rain Forests

Shogoro Fujiki 1,*, Ryota Aoyagi 1, Atsushi Tanaka 1,2, Nobuo Imai 1,3, Arif Data Kusma 4, Yuyun Kurniawan 5, Ying Fah Lee 6, John Baptist Sugau 6, Joan T. Pereira 6, Hiromitsu Samejima 7 and Kanehiro Kitayama 1

1 Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan; aoyagi.ryota@gmail.com (R.A.); amcw2003@gmail.com (A.T.); i96nobuo@gmail.com (N.I.); kanehiro@kais.kyoto-u.ac.jp (K.K.)
2 Japan Forest Technology Association, Tokyo, Japan
3 Primate Research Institute, Kyoto University, Aichi, Japan
4 WWF Indonesia, Kutai Barat, Kalimantan Timur, Indonesia; AKusuma@wwf.or.id
5 WWF Indonesia, Ujung Kulon, Banten, Indonesia; YKurniawan@wwf.or.id
6 Forest Research Centre, Sabah Forestry Department, 90715 Sandakan, Sabah, Malaysia; yingfah.lee@sabah.gov.my (Y.F.L.); john.sugau@sabah.gov.my (J.B.S.); Joan.pereira@sabah.gov.my (J.T.P.)
7 Institute for Global Environmental Strategies, Kanagawa, Japan; lahang.lejau@gmail.com

* Correspondence; fujiki5636@gmail.com; Tel.: +81-75-753-6080.

Academic Editors: Jeffrey Sayer and Chris Margules
Received: 27 July 2016; Accepted: 6 December 2016; Published: 11 December 2016

Abstract: Assessment of the progress of the Aichi Biodiversity Targets set by the Convention on Biological Diversity (CBD) and the safeguarding of ecosystems from the perverse negative impacts caused by Reducing Emissions from Deforestation and Forest Degradation Plus (REDD+) requires the development of spatiotemporally robust and sensitive indicators of biodiversity and ecosystem health. Recently, it has been proposed that tree-community composition based on count-plot surveys could serve as a robust, sensitive, and cost-effective indicator for forest intactness in Bornean logged-over rain forests. In this study, we developed an algorithm to map tree-community composition across the entire landscape based on Landsat imagery. We targeted six forest management units (FMUs), each of which ranged from 50,000 to 100,000 ha in area, covering a broad geographic range spanning the most area of Borneo. Approximately fifty 20 m-radius circular plots were established in each FMU, and the differences in tree-community composition at a genus level among plots were examined for trees with diameter at breast height ≥10 cm using an ordination with non-metric multidimensional scaling (nMDS). Subsequently, we developed a linear regression model based on Landsat metrics (e.g., reflectance value, vegetation indices and textures) to explain the nMDS axis-1 scores of the plots, and extrapolated the model to the landscape to establish a tree-community composition map in each FMU. The adjusted R² values based on a cross-validation approach between the predicted and observed nMDS axis-1 scores indicated a close correlation, ranging from 0.54 to 0.69. Histograms of the frequency distributions of extrapolated nMDS axis-1 scores were derived from each map and used to quantitatively diagnose the forest intactness of the FMUs. Our study indicated that tree-community composition, which was reported as a robust indicator of forest intactness, could be mapped at a landscape level to quantitatively assess the spatial patterns of intactness in Bornean rain forests. Our approach can be used for large-scale assessments of tree diversity and forest intactness to monitor both the progress of Aichi Biodiversity Targets and the effectiveness of REDD+ biodiversity safeguards in production forests in the tropics.
Keywords: Aichi Biodiversity Targets; environmental safeguards; forest intactness; logged-over forests; REDD+; satellite remote sensing; tropical production forests

1. Introduction

Continued deforestation and forest degradation and the associated losses of biodiversity in tropical countries represent major global concerns [1]. To date, coordinated international efforts have resulted in two international conventions that attempt to reduce the rate of tropical deforestation and forest degradation and the associated biodiversity losses: the Convention on Biological Diversity (CBD), and the United Nations Framework Convention on Climate Change (UNFCCC).

The CBD sets a strategic plan for biodiversity for 2011–2020 and the Aichi Biodiversity Targets, which include several targets for forest conservation and their sustainable use [2]. It is important to quantitatively assess the progress of the Aichi Biodiversity Targets, and there has been intensive discussion on which indicators should be used for monitoring them [3,4]. The development of essential biodiversity variables as a measurement for studying, reporting, and managing changes in biodiversity is a prerequisite for achieving the targets [4].

The UNFCCC is a convention primarily targeting the mitigation of and adaptation to climate change. Reducing Emissions from Deforestation and Forest Degradation Plus (REDD+), a post-Kyoto Protocol mechanism developed under UNFCCC, is expected to have positive effects on biodiversity conservation because its main targets are natural tropical forests. When considering the impact of REDD+, it is important to consider REDD+ activities with biodiversity trade-offs. For example, if low-carbon, high-biodiversity tropical forests are converted into high-carbon, low-biodiversity forests (e.g., plantations and forests dominated by pioneer species), REDD+ may have an overall negative effect on biodiversity conservation [5,6]. Measures to safeguard biodiversity have been extensively discussed in this context under UNFCCC [7–9]. The premise of these safeguards is a compliance system where each REDD+ project is required to comply with standards and indicators to qualify for REDD+ credits. Such a compliance system must involve third party auditing and verification based on standards and indicators. Therefore, there is an urgent need to develop robust indicators to both accurately assess the progress of the Aichi Biodiversity Targets and avoid overall negative effects on biodiversity caused by the implementation of REDD+. A key obstacle to achieving these goals is the lack of global, integrated observation systems for delivering regular, timely data on changes in biodiversity [4].

Indicators for compliance systems can be generic or specific. Generic indicators [10] are a set of indicators applicable to all forest types and regions; they may be largely based on readily available statistical data or actual prescribed management plans such as the presence/area of forests with high conservation value, presence/type of conservation measures, and presence/magnitude of mitigation measures. Specific indicators are a set of indicators that involve direct measurements in the field such as changes in the distribution/number of endangered species, changes in the area of intact ecosystems, and changes in species richness. For indicators to be reliable and practical, they must incur a low financial cost, be easily identified, be proxies for ecosystem integrity, and have cross-taxon congruency [11] as well as sufficient robustness and sensitivity [12]. Generic indicators are superior in cost effectiveness, while specific indicators are superior as proxies of ecosystem integrity and sensitivity.

Recently, Imai et al. [12] proposed that the community composition of canopy trees could be used as a specific indicator for “ecosystem integrity or intactness” in spatiotemporally dynamic Bornean production forests where timber is produced by commercial logging. The authors used the axis-1 scores of the ordination of vegetation plots based on the relative species (or genus) abundances in logged-over forests to indicate forest intactness. Derived axis-1 scores showed significant linear correlations with magnitude of logging intensity (i.e., the inverse of remaining above-ground biomass). This significant linear correlation was a result of the interaction of linearly
increasing species number in the pioneer guild and linearly decreasing species number in the climax guild with increased logging intensity [12,13]. Therefore, the indicator based on community composition (i.e., axis-1 scores of the ordination) actually indicated the compositional distance from an intact forest where there were zero or minimal effects from logging. The count-plot measurements on the ground were fast and inexpensive, and implemented by local foresters [12]. Therefore, the methods of Imai et al. [12] satisfied the above-listed requirements for indicators.

A previously unresolved issue related to specific indicators is spatial representativeness; specific indicators derived from plots on the ground need to be extrapolated to the landscape or even region level. To address the increasing need for a practical indicator that is spatially/temporally sensitive and robust, we developed a new algorithm using satellite remote sensing to map forest intactness in Bornean tropical rain forests based on the community composition indicator of Imai et al. [12]. We used generic abundances instead of specific abundances because both specific and generic abundances demonstrated the same response to logging intensity [12]. Specifically, we developed a new algorithm to map variation in tree-community composition using Landsat imagery as a proxy for forest intactness and assessed whether our method could reliably diagnose the magnitude of forest intactness among/within production forests under different management regimes. We developed a linear regression model based on Landsat metrics to explain the community composition indicator of Imai et al. [12] and extrapolated the model to the landscape to establish a tree-community composition map in each forest management unit (FMU). Such a method will provide a valuable contribution to assessments of the progress of meeting the Aichi Biodiversity Targets and effectiveness of REDD+ safeguards as well as improved management of forest management units (FMUs).

2. Materials and Methods

2.1. Study Site

We studied the community composition of canopy trees and mapped the patterns of tree-community composition in six Bornean FMUs that were conducting legal selective logging for commercial purposes. An FMU is a management entity with a valid logging license from the local government either from Malaysian states in Malaysian Borneo or from the Indonesian Government in Indonesian Borneo. In each FMU, the marketplace trees (large dipterocarp species) within a certain size class are harvested for timber. FMUs in Borneo typically range from 50,000 to 100,000 ha in size, and their land belongs to the respective governments. These areas are covered by lowland mixed dipterocarp forests, with varying degrees of degradation reflecting their different logging histories.

The studied FMUs were Segaliud Lokan (5°20′–27′N, 117°23′–39′E, 576 km²), Deramakot (5°14′–28′N, 117°20′–38′E, 551 km²), Tangkulap (5°18′–31′N, 117°11′–22′E, 276 km²), and Sapulut (4°40′–55′N, 116°30′–117°00′E, 956 km²) in Sabah, Malaysia; and, Roda Mas (0°46′–1°05′N, 114°25′–115°06′E, 703 km²) and Ratah (0°7′S–0°13′N, 114°58′–115°30′E, 982 km²) in East Kalimantan, Indonesia (Figure 1). The areas of Segaliud Lokan, Deramakot, and Tangkulap were initially logged in 1958, 1956, and 1970, respectively, using conventional logging methods (i.e., high-impact logging with no environmental considerations [14]). The three FMUs were adjacent to one another. In Deramakot, conventional logging continued until 1989, when all logging activities were halted for regrowth. Then, a long-term management plan with reduced-impact logging was introduced to Deramakot in 1995. Reduced-impact logging is an improved method of selective logging, including pre-harvest inventory, mapping of all canopy trees, directional felling, liana cutting, and planning of skid trails, log decks, and roads [15,16]. In combination with reduced-impact logging, a longer cutting cycle (i.e., 40 years) was strictly adhered to in accordance with the long-term management plan [17]. These combined approaches helped to preserve forest integrity [13,18–20]. Deramakot was the first tropical forest certified by the Forest Stewardship Council (FSC) in 1997, and was considered an exemplary model of sustainable forest management by the Sabah Government. Reflecting this logging history, we observed that the forests inside the Deramakot FMU were less
disturbed based on aboveground biomass and varied biological communities [13]. In contrast, Segaliud Lokan was repeatedly logged using conventional logging until 2002, after which reduced-impact logging has been implemented. Tangkulap was repeatedly logged using conventional logging until 2002, after which all logging activities have remained suspended. Sapulut was first logged in 1956 and repeatedly logged until 2000 using conventional logging, after which reduced-impact logging has been carried out, but industrial tree plantations have been established and Sapulut now consists of rubber and Acacia plantations over approximately 35% of its area. Reduced-impact logging was implemented in Roda Mas from at least 2008. It is highly likely that logging had been conducted earlier in the Roda Mas area, but there is no information about its logging history before 2008. We estimate that the logging intensity was relatively mild in the Roda Mas area because areas of intact forest still remain throughout the area. Ratah implemented conventional logging from 1972 to 2010, after which reduced-impact logging was implemented. Deramakot, Tangkulap, Roda Mas, and Ratah are all certified by the FSC. Sapulut and Segaliud Lokan are certified by Malaysian Timber Certification Council (MTCC). Table 1 summarizes the target areas.

Figure 1. Locations of the six forest management units (FMUs) in this study: (1) Segaliud Lokan, (2) Deramakot, (3) Tangkulap, (4) Sapulut, (5) Roda Mas, and (6) Ratah.

Table 1. Profiles of the forest management units (FMUs) targeted in this study.

<table>
<thead>
<tr>
<th>Name of FMU</th>
<th>State/Province, Nation</th>
<th>Silviculture and Timber Harvest</th>
<th>Forest Certification</th>
<th>Established Plot Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segaliud Lokan</td>
<td>Sabah, Malaysia</td>
<td>1958–2002 CL, Since 2003 RIL</td>
<td>MTCC</td>
<td></td>
</tr>
<tr>
<td>Tangkulap</td>
<td></td>
<td>1970–2002 CL</td>
<td>FSC</td>
<td></td>
</tr>
<tr>
<td>Sapulut</td>
<td></td>
<td>1956–2000 CL, Since 2001 RIL</td>
<td>MTCC</td>
<td>50</td>
</tr>
<tr>
<td>Roda Mas</td>
<td>East Kalimantan,</td>
<td>Since at least 2008 RIL</td>
<td>FSC</td>
<td>50</td>
</tr>
<tr>
<td>Ratah</td>
<td>Indonesia</td>
<td>1972–2010 CL, Since 2011 RIL</td>
<td>FSC</td>
<td>90</td>
</tr>
</tbody>
</table>

Note: CL, conventional logging; RIL, reduced impact logging; FSC, Forest Stewardship Council; MTCC, Malaysian Timber Certification Council.
2.2. Field Survey

We conducted count-plot surveys in the four FMUs: Segaliu Lokan, Sapulut, and Ratah in 2011–2012 [12] and Roda Mas and Ratah in 2014–2015. The number of the established plots in each FMU is shown in Table 1. Because Segaliu Lokan, Deramakot, and Tangkulap are adjacent to one another with the same forest type, we assumed that the count-plot data obtained from the field survey in Segaliu Lokan was representative of the canopy species composition of the forests in all three FMUs.

Details of the procedure of the count-plot survey are described in Imai et al. [12]. Briefly, to select representative vegetation plots from a heterogeneous forest with varying magnitudes of forest degradation, we classified each FMU into five strata based on Landsat imagery according to the extent of forest degradation, from intact forest (stratum 1) to open canopy area with several pioneer trees (stratum 5) [12]. Here, we defined the magnitude of forest degradation based on the remaining aboveground biomass, a definition of the Intergovernmental Panel on Climate Change [21]. Ten 20 m radius circular plots (1257 m2 in area) were randomly established at altitudes below 600 m a.s.l. in each stratum in each FMU (i.e., 50 plots in each FMU). To minimize spatial autocorrelation, plots were established at a distance of at least 100 m from one another. Global positioning system (GPS) data were collected during the field surveys at the center of a plot with calibration for at least one hour in each plot. We measured diameter at breast height (dbh) of all trees (dbh ≥10 cm) and counted all measured species in each plot. Woody vines were excluded from the inventory. Trees with buttresses were measured well above (ca. 50 cm) protrusions. All trees were identified by local botanical experts. If trees could not be identified in the field, voucher specimens were collected and identified in local herbaria. Samples that could not be identified to the species level were distinguished as morphospecies. We obtained inventory data from a total of 240 plots.

2.3. Field Data Analysis

We recorded 14,783 stems from 85 families and 254 genera in 240 plots (totaling 30.2 ha in area). The Chao distances [22] and the number of trees of each genus were used to calculate the distance matrix for the inventoried plots in each FMU. Then, an ordination of plots was conducted to ordinate the inventoried plots (approximately 50 plots) for each FMU with non-metric multidimensional scaling (nMDS) using the metaMDS procedure in the vegan package in the R software program (R version 3.2.0) [23]. Subsequently, the nMDS axis-1 scores were used as an index of tree-community composition. Imai et al. [12] demonstrated that the nMDS axis-1 scores of plots correlated with the aboveground biomass values of the plots, a proxy for forest degradation, in each FMU (R²=0.52–0.71). Although plots were a considerable distance from one another, we were unable to completely rule out the possibility of spatial autocorrelations among plots. Therefore, we tested for autocorrelations among plots in each FMU but found none (see Supplementary Materials, Figures S1 and S2). Hence, the nMDS axis-1 scores of plots indicated intactness in terms of tree-community composition—in other words, the compositional distance from intact forests where there were no effects of logging. Intactness can also be functionally translated to ecosystem integrity because regeneration ability is assured in more intact forests.

To compare the differences between FMUs with the same index, the nMDS axis-1 scores were normalized by transforming the scores into new scores with a mean of 0 and a standard deviation of 1. In order to assure that a given normalized score can indicate the same forest condition in terms of forest intactness across FMUs, we regressed normalized nMDS axis-1 scores with relative abundance of pioneer species (genera) per plot in each FMU. Subsequently, significant differences in the slope and intercept of the regression lines among FMUs were tested with a multi regression analysis. Complete overlap of the regression lines suggests that a given normalized score can indicate the same forest condition across FMUs. Indeed, there were no significant differences in the intercept and slope among the regression lines in Segaliu Lokan, Ratah and Sapulut (Supplementary Materials, Figure S3). The intercept and slope of Roda Mas were exceptionally
significantly different from the other FMUs, but the range of the deviation was still within that of
the other FMUs. The normalized scores will be used as “nMDS axis-1” in the following analyses.

2.4. Satellite Analysis

2.4.1. Satellite Images and Image Pre-Processing

The estimates of the nMDS axis-1 scores for the entire study area in each FMU were based on
Landsat TM (Thematic Mapper) and OLI (Operational Land Imager) imageries. Descriptions of the
data set (sensor, path/row and date) of the imageries used are given in the Supplementary Material
(Table S1). No significant geolocation errors were observed in the images because the GPS ground
positions collected on logging roads corresponded closely with the images.

As pre-processing, the raw digital numbers of each image were converted into
top-of-atmosphere radiances. Subsequently, to compensate for atmospheric scattering and
absorption effects, an atmospheric correction algorithm based on the Second Simulation of a
Satellite Signal in the Solar Spectrum radiative transfer code [24,25] was used to convert
top-of-atmosphere radiances into surface reflectance. Finally, the effects of differential illumination
due to topography were reduced using the method described by Ekstrand [26]. Shuttle Radar
Topography Mission (SRTM) data were used to correct the illumination effects. Then, pixels
covered with clouds/shadow were removed using an object-based approach. We created image
segments composed of spectrally coherent pixels that were clustered based on homogeneity criteria,
and established a threshold to detect cloud/shadow segments. Then, we removed them using the
threshold and visual inspection.

Subsequently, the missing data due to cloud cover in Sapulut, Roda Mas, and Ratah were filled
in using the cloud-free areas of temporally adjoining data. Where satellite imagery is subjected to
high incidence of clouds or haze, which is often the case in tropical rain forests, the mosaicking of
cloud-free parts of temporally adjoining data may be the best option for deriving cloud-free
coverage [27]. In our study, in the mosaicking procedure, calibrated cloud-free parts of adjoining
secondary images were embedded in the missing parts of the base image. To calibrate secondary
images to the base image, we used a linear regression model [28–34]. The basis for the model was
the set of co-located, mutually clear pixels from each base- and secondary-scene pair. Before
establishing regression models, we excluded pixels with marked spectral changes, which possibly
occurred with land-cover changes, using a change-detection analysis with a normalized difference
vegetation index (NDVI) [35]. Subsequently, pixel-level regression models were established using a
robust regression. A regression model for Landsat-TM/OLI images has the following general form:

\[ y_{\text{base}} = f (x_{\text{sec}1}, x_{\text{sec}2}, x_{\text{sec}3}, x_{\text{sec}4}, x_{\text{sec}5}, x_{\text{sec}6}, x_{\text{sec}7}), \]

where \( y_{\text{base}} \) is the reflectance value of the pixel in the base image for the \( i \)th band to be predicted,
and \( x_{\text{sec}1} \) is the reflectance value of band 1 of the co-located pixel in the secondary images, \( x_{\text{sec}2} \) is the
reflectance value of band 2, and so on. Missing data in each pixel in the base images were replaced
by the calibrated pixels of secondary images using the regression models. Landsat-OLI
(Jun/19/2013), Landsat-OLI (May/31/2015), and Landsat-TM (Feb/10/2010) were used as base images
in Sapulut, Roda Mas, and Ratah, respectively. eCognition Developer 8.7 was used to create image
segments, and ERDAS Imagine ver.11.0 and ArcGIS 9.3.1 were used for pre-processing.

2.4.2. Extrapolation of nMDS Axis-1 Scores Based on Landsat Data

To estimate the nMDS axis-1 scores of the entire study area of each FMU, we established
multiple linear regression models between the nMDS axis-1 scores of the inventoried plots and
the corresponding metrics of the Landsat imagery (i.e., average value of Landsat metrics within the 20
m-radius from the center of the plots) in each FMU. Then, the nMDS axis-1 scores were
extrapolated to the entire area of each FMU based on the model. The inventory plots that were
unavailable because of cloud cover and could not be replaced in the above cloud-correction
procedure were eliminated in this procedure. We developed one independent model for each FMU
rather than developing a single model for all FMUs combined by combining all plots because it was necessary to take regional floristic variation into account.

The normalized scores of nMDS axis-1 were considered response variables, and the Landsat metrics were considered independent variables. We used the following Landsat metrics: reflectance value of each band (Band1\_TMOLI, Band2\_TMOLI, Band3\_TMOLI, Band4\_TMOLI, Band5\_TMOLI, Band6\_OLI, Band7\_TMOLI), normalized difference vegetation index (NDVI) [35], normalized difference water index (NDWI) [36,37], normalized difference soil index (NDSI) [38], and enhanced vegetation index (EVI) [39]. The mean values of the Landsat metrics within a 20 m radius from the center of the plots were used as the metrics corresponding to each plot. The indices were calculated from the following equations:

\[
\text{NDVI} = \frac{(\text{Band4}\_\text{OLI} - \text{Band3}\_\text{OLI})}{(\text{Band4}\_\text{OLI} + \text{Band3}\_\text{OLI})},
\]
\[
\text{NDWI} = \frac{(\text{Band3}\_\text{OLI} - \text{Band5}\_\text{OLI})}{(\text{Band3}\_\text{OLI} + \text{Band5}\_\text{OLI})},
\]
\[
\text{NDSI} = \frac{(\text{Band5}\_\text{OLI} - \text{Band4}\_\text{OLI})}{(\text{Band5}\_\text{OLI} + \text{Band4}\_\text{OLI})},
\]
\[
\text{EVI} = 2.5 \cdot \frac{(\text{Band4}\_\text{OLI} - \text{Band3}\_\text{OLI})}{(\text{Band4}\_\text{OLI} + 6 \cdot \text{Band3}\_\text{OLI} - 7.5 \cdot \text{Band1}\_\text{OLI} + 1)}.
\]

In addition to the above metrics, the coefficient of variation (CV), standard deviation (SD), and textures of the gray level co-occurrence matrix (GLCM) [40] were used as proxies for spectral heterogeneity because degradation of forest canopies might affect the heterogeneity of the spectral pattern. The CV, SD, and textures of the GLCM were calculated using a 3×3 pixel window based on the reflectance values and each of the indices. The GLCM is a tabulation of how often different combinations of gray levels occur at a specified distance and orientation in an image object [41]. The homogeneity, contrast, angular second moment, entropy, dissimilarity, correlation, mean, and standard deviation were calculated as the indices of the textures. A total of 120 and 132 metrics were generated based on Landsat TM and OLI, respectively. The independent variables of the regression models were chosen using a stepwise selection from all of the metrics to avoid multi-collinearity among the independent variables. The values of the variance inflation factor (VIF) of selected independent variables were all less than 10 except for Roda Mas, which indicated that there was no multicollinearity (see Supplementary Materials, Table S2). We removed the areas above 600 m elevation from the analysis, because the potential natural vegetation above 600 m differed from that of lowland natural forest [42,43]. eCognition developer 8.7 was used to calculate texture, R ver. 3.20 was used to establish the models, and ERDAS Imagine ver.11.0 and ArcGIS 9.3.1 were used in the other procedures.

2.5. Validation of nMDS Axis-1 Model

We used a cross-validation approach [20,44] to assess the accuracy of the models. Four-fifths of all plots were randomly selected in each FMU to construct the nMDS axis-1 models. Based on these models, we estimated the nMDS axis-1 scores of the remaining one-fifth of the plots. Then, we tested the correlation between the estimated nMDS axis-1 scores and the field-measured nMDS axis-1 scores. This step was reiterated 1,000 times in each FMU to derive the 95% confidence interval of the correlation coefficient. The statistical tests were conducted using R ver. 3.20.

2.6. Comparison of Canopy Conditions Based on nMDS Axis-1 Scores among FMUs

To assess whether our method could reliably diagnose the magnitude of forest intactness, we calculated the histogram of relative frequency and mean value of the nMDS axis-1 scores for each FMU and compared them based on their management types. To compare the histograms, we chose four representative FMUs: Roda Mas (FMU practicing reduced-impact logging in relatively intact forests); Deramakot (FMU practicing reduced-impact logging in previously conventionally logged forests); Segaliud Lokan (FMU with a longer history of conventional logging); and Sapulut (FMU including industrial tree plantations). The histograms were represented as probability density functions, which indicate relative frequency of nMDS axis-1 scores in each FMU.

The histogram of the nMDS axis-1 scores of each FMU was used as an index of the spatial variability of the magnitude of forest intactness, and the mean value was used as the average condition in each FMU. To compare the histograms of the nMDS axis-1 scores among FMUs, a
regression model based on all plots (approximately 50) was extrapolated to the entire area of the FMU, from which the histogram was derived. To compare the differences in the mean nMDS axis-1 scores among the six FMUs, four-fifths of all plots were randomly selected to construct nMDS axis-1 models in each FMU. Based on these models, we estimated the mean nMDS axis-1 scores of the entire area of each FMU. This step was iteratively done 1000 times to derive the mean value and 95% confidence intervals of the mean nMDS axis-1 scores. Then, the relationships of the derived histograms and means of the nMDS axis-1 scores with the type of forest management were examined.

3. Results

3.1. Forest-Intactness Maps

The nMDS axis-1 scores based on the stepwise selection were mainly explained by the short-wave infrared (SWIR) reflectance, textures, and SD (Table 2). The correlation coefficient scores between the predicted and observed nMDS axis-1 scores ranged from 0.604 to 0.745 (Figure 2). The derived forest-intactness maps of all FMUs are shown in Figure 3. Each pixel in the maps contains a nMDS axis-1 score, and the color gradation from blue to red indicates a gradient of nMDS axis-1 scores from high to low values as a proxy for the canopy tree composition at a genus level. The maps showed conspicuous blue areas corresponding to the tree community composition of an intact forest as well as conspicuous red areas corresponding to the tree-community composition of the least intact forest. Intact forests and least-intact forests were indicated as the two extremes of a continuum of forest degradation in our method.

| Table 2. Description of established multivariate regression models based on all plots. |
|----------------------------------|-----------------|-----------------|-----------|--------|
|                                   | R²              | Coefficient     | SE        | T-value | Pr (>|t|) |
| Segaliud-Lokan-Deramakot-Tangkulap (N=47) |                 |                 |           |        |
| B7_TM                            | 0.642           | 0.38989         | 2.80E-04  | 3.69863 | <0.001 |
| GLCM_mean_NDVI                   | -0.25902        | 1.20E-02        | 2.23543   | 3.10E-02|
| GLCM_correlation_B3_TM           | 0.32586         | 3.80E-01        | 3.3816    | 1.60E-03|
| GLCM_mean_B1_TM                  | 0.30875         | 9.40E-03        | 2.76426   | 8.50E-03|
| GLCM_homogeneity_B4_TM           | 0.1891          | 3.70E+00        | 2.03608   | 4.80E-02|
| Sapulut (N = 45)                 |                 |                 |           |        |
| B6_OLI                           | -0.46159        | 2.40E-05        | 3.7795    | <0.001 |
| GLCM_contrast_B3_OLI             | 0.37483         | 3.10E-05        | 3.8679    | <0.001 |
| B5_OLI                           | -0.36385        | 1.20E-05        | 3.0253    | 4.30E-03|
| GLCM_mean_B6_OLI                 | 0.286           | 8.20E-03        | 2.8178    | 7.50E-03|
| Roda Mas (N = 45)                |                 |                 |           |        |
| B6_OLI                           | -0.86698        | 7.50E-05        | 4.9699    | <0.001 |
| SD_NDSI                          | -0.75056        | 1.20E-04        | 4.5416    | <0.001 |
| SD_NDWI                          | 0.745           | -2.12925        | 2.20E-04  | 4.6781  | <0.001 |
| SD_NDVI                          | 2.00116         | 4.50E-04        | 3.373     | 1.70E-03|
| B4_OLI                           | 0.56146         | 4.70E-04        | 2.3481    | 2.40E-02|
| Ratah (N = 64)                   |                 |                 |           |        |
| B7_TM                            | -1.42357        | 1.30E-03        | 7.02905   | <0.001 |
| B3_TM                            | 0.615           | 0.75859         | 2.80E-03  | 3.74366 | <0.001 |
| GLCM_dissimilarity_B4_TM         | 0.25656         | 9.80E-03        | 3.2506    | 1.90E-03|

Note: N, number of plots used for each model; R², adjusted R-squared value; Coefficient, standardized partial regression coefficient; SE, standard error; SD, standard deviation; GLCM, grey level co-occurrence matrix; NDVI, normalized difference vegetation index; NDSI, normalized difference soil index; NDWI, normalized difference water index.
Figure 2. Model fits for all available inventory plots in each FMU. Scatter plots show relationships between predicted and observed scores.

Figure 3. Tree-community composition maps of all FMUs. A color gradation from blue to red indicates a gradient of normalized nMDS axis-1 scores from high to low values as a proxy for canopy tree composition at a genus level.

Several FMUs had areas of intact forests (e.g., along the left side in Ratah and upper-right side in Roda Mas). A few of the FMUs had conspicuous red-yellow areas, which corresponded to highly degraded areas. The most conspicuous red-yellow area was the eastern part of Ratah, which corresponded to the land after forest fire in 1998. The red-yellow area at western part of Roda Mas was allocated to local people who practice slash and burn agriculture. Mosaics of red areas occurred in Sapulut, which corresponded to industrial tree plantations. The forest intactness inside the
Deramakot was uniformly green, reflecting the history of reduced-impact logging and longer cutting cycle (i.e., 40 years). In contrast, Segaliud Lokan where reduced-impact logging had been implemented after 2002 and Tangkulap where all logging activities had remained suspended after 2002 had sparse yellow-red areas inside. This indicated that the negative effect of high-impact conventional logging continued until 2002 still remained in these two FMUs.

3.2. Validation of Community-Composition (nMDS Axis-1) Models

The cross validation indicated that the accuracy of the model estimates varied among FMUs but was generally high. The means (95% confidence interval, CI) of the iteratively calculated correlation coefficients between the estimated nMDS axis-1 scores and actual observed scores (adjusted $R^2$ values) were 0.57 (CI 0.10–0.87), 0.54 (CI 0.10–0.87), 0.69 (CI 0.37–0.91), and 0.56 (CI 0.12–0.84) for Segaliud Lokan-Deramakot-Tangkulap, Sapulut, Roda Mas, and Ratah, respectively.

3.3. Comparison of Histogram/ Mean nMDS Axis-1 Scores among FMUs

The histograms of the nMDS axis-1 scores in six FMUs are shown in Figure 4a. The histograms varied significantly among the FMUs in terms of mode and pattern; the ranges were the same because the nMDS axis-1 scores were standardized across the FMUs. The mode of Roda Mas had the highest axis-1 score among the FMUs, indicating that relatively intact forests occur over a disproportionately greater area in Roda Mas. At the same time, the density of low nMDS axis-1 scores (nMDS axis-1 score = −1) was also nominal in Roda Mas, indicating the occurrence of current logging activities with reduced-impact logging techniques. The mode in Deramakot, the model site of sustainable forest management in Sabah, had the second highest score among the FMUs. The nMDS axis-1 scores of the modes among four representative FMUs were decreased in the order reflecting their logging intensity, history or management (Figure 4b).

![Figure 4. Histograms of normalized nMDS (non metric multidimensional scaling) axis-1 scores (a) in all FMUs, and (b) in four selected FMUs. (RIL, reduced-impact logging; CL, conventional logging).](image)

The mean values and 95% confidence intervals of the nMDS axis-1 scores are shown in Figure 5, in which the six FMUs are arranged in increasing order of their mean nMDS axis-1 score. Deramakot, which has been designated as a model of sustainable forest management by the Sabah Government, recorded the highest mean score, 0.318 (CI 0.197–0.432). The second-highest score was 0.289 (CI 0.219–0.365) in Roda Mas. The scores were 0.220 (CI 0.151–0.293) in Ratah. Tangkulap and Segaliud Lokan, where high-impact conventional logging continued until 2002, had relatively low scores of 0.192 (CI 0.094–0.290) and 0.163 (CI 0.063–0.262). Sapulut, where industrial tree plantations occur in approximately 35% of the area, recorded the lowest score of -0.082 (CI -0.184–0.027).
Overall, the mean score of Sapulut (-0.082) was significantly lower than that of the other FMUs, while the other FMUs did not significantly differ from one another. There was well-defined variation in the mean score (0.192–0.318) among the FMUs that were certified by the FSC (Figure 5), although the 95% confidence intervals broadly overlapped with each other among the FMUs.

Figure 5. Mean values of normalized nMDS axis-1 scores in all FMUs.

4. Discussion

To address the increasing need for a practical method of assessing biodiversity that is spatially and temporally sensitive and robust, we developed an algorithm to map forest intactness based on tree-community composition. As a proxy for tree-community composition, we used the nMDS axis-1 score, a robust and sensitive indicator of forest intactness in Bornean natural production forests [12]. The maps based on the algorithm demonstrated a significant variation in composition of tree genera both among and within FMUs.

To date, several remote sensing methods have been developed to map tree-community composition using imaging spectroscopy [45-49]. Where imaging spectroscopy (also known as hyperspectral imagery) can characterize variation in tree compositional assemblages, multispectral imageries (e.g., Landsat) have been considered to lack the spectral resolution to effectively demonstrate compositional variation [50,51]. However, our study shows that multispectral remote-sensing data of Landsat images are potentially useful for characterizing the variation in composition of tree genera caused by logging and forest degradation. The concept of mapping compositional gradients using remote sensing is based on the assumption that the compositional gradients of vegetation are related to variation in canopy traits such as reflectance spectra, foliar chemistry, morphological traits, canopy form, and canopy structure [48]. Clearly, Landsat imagery may not effectively resolve subtle intra-guild trait variation (e.g., trait variation within climax guild) because of the low spatial/spectral resolution; however, Landsat imagery has been proven to be capable of detecting inter-guild trait differences (i.e., trait differences between pioneer and climax guilds) [52,53] as shown by our algorithm. The reason why the proxy for community composition (nMDS axis-1 score) used in our study is able to reflect canopy intactness is that the nMDS axis-1 scores can reflect the interaction of linearly increasing pioneer guild and linearly decreasing climax guild with increasing logging intensity [12]. Thus, the selected independent variables in our models and Landsat imageries reflect the interactions of these guilds. The variables selected by stepwise selection were principally related to SWIR reflectance and spectral heterogeneity information (i.e., texture and SD). SWIR reflectance is significantly related to stand age, height, volume, and biomass in tropical secondary forest [54]. In contrast, spectral heterogeneity information can be considered a
proxy for species diversity [55]. These patterns were also apparent in our study. SWIR reflectance was a proxy for tree growth (i.e., biomass) after logging, while texture and SD were related to the heterogeneity of the two major regeneration guilds. We suggest that SWIR reflectance and spectral heterogeneity information can reflect the interactions of the two guilds, and our algorithm is thus suitable for the characterization of the variation in composition of tree genera caused by logging disturbances.

It should also be noted that the selected independent variables and coefficients varied among models. Our study area covered the entire area of Borneo, spanning a large geographic area, and included significant floristic differences among the tropical rain forests [56]. Therefore, the selected independent variables would be expected to vary among FMUs, reflecting variation in canopy structure and foliar traits.

The mean adjusted \( R^2 \) values (0.54–0.69) based on the cross-validation approach showed close correlation in all FMUs. The mean adjusted \( R^2 \) values were high despite the use of moderate-resolution Landsat data and the semi-qualitative nature of our metrics (i.e., tree-community composition), which were based on the mixing ratio of the canopy genera. However, the range of adjusted \( R^2 \) values (0.1–0.9) at the 95% confidence interval across all FMUs was fairly wide; this probably indicates that the number of plots used for cross validation was too small. Depending on the combination of plots used for developing the models, the correlation between the predicted and observed nMDS axis-1 scores became very low. Because of the wide variation in the \( R^2 \) values at the 95% confidence interval, it was difficult to discriminate the FMUs from one another, despite the fact that the mean nMDS axis-1 scores per se differed considerably among FMUs (Figure 5). Overall, we were able to differentiate only one FMU with the lowest mean nMDS axis-1 score (Sapulut) from the other FMUs (Figure 5) in our analysis. However, our method is clearly useful for differentiating between intact and highly disturbed forests within each FMU when all 50 plots are used to develop the best model. It can identify the location and extent of intact forests as areas of potentially high conservation value, or those of least-intact forests where mitigation measures are required. In combination with a biomass estimate, our method is also useful for identifying the location and extent of less-intact forests with high biomass, indicative of plantations. It thus offers a potential tool for verifying compliance with environmental safeguards in REDD+, in which a primary concern is the conversion of natural forests into plantation.

The applicability of our method was verified over a broad geographic range (0°7’S–5°20’N, 114°25’–115°30’E) spanning the most area of Borneo. The principle of using tree-community composition may also be applicable to other forests outside Borneo because the occurrence of the two major regeneration guilds (i.e., the pioneer and climax guilds) is a common biological phenomenon; however, this needs to be tested in the future. One of the advantages of this method is the broad coverage and low cost based on Landsat imagery. Moreover, our analysis of community composition is based on canopy-tree genera only. As Imai et al. [12] pointed out, identification of canopy trees at the generic level can reduce the cost and time of identification by at least 60% compared with the conventional routine survey at species level. The spatial and temporal availability of Landsat imagery is superior to other remote sensing data, and therefore the Landsat-based tree-community composition map is suitable for regional and global biodiversity monitoring. These advantages are essential for a global biodiversity assessment of the Aichi Biodiversity Targets and REDD+ biodiversity safeguards. We propose a practical method combining count-plot surveys on the ground with Landsat remote sensing for large-scale forest biodiversity/ecosystem assessments of both the Aichi targets and REDD+ biodiversity safeguards in natural forests in the tropics.

Acknowledgements: We acknowledge the kind assistance from WWF (World Wide Fund for Nature) Indonesia, WWF Japan, the Agriculture and Forestry Services of the District of Mahakam Ulu (East Kalimantan), Sabah Forestry Department and Sabah Forest Research Centre, Sarawak Forest Department, Botanical Research Center, Sarawak Forestry Corporation, the KTS Plantation Sdn. Bhd., the Sapulut Forest Development Sdn. Bhd., Zedtee Sdn. Bhd., the PT (Perseroan Terbatas) Roda Mas Timber and the PT Ratah Timber. We are thankful to Miun Poster, Joel Bin Dawat and Sutrisno for tree species identification; Itong Sarjuni and the staff of the FMUs for assisting field inventory; and Dodit Agus Rijono, and Mutai Hashimoto
for generous support for every aspect. This study was supported by the Global Environment Research Fund 1-1403 of the Ministry of the Environment, Japan, to Kanehiro Kitayama.

Author Contributions: K.K. conceived and designed the research; S.F. invented the algorithm; S.F., R.A., A.T., N.I. and H.S. collected and analyzed the data; and all authors wrote the paper.

Conflicts of Interest: The authors declare that there is no conflict of interests regarding the publication of this paper. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References


20. Langner, A.; Samejima, H.; Ong, R.C.; Titin, J.; Kitayama, K. Integration of carbon conservation into


