Towards nationwide mapping of bamboo resources in the Philippines: testing the pixel-based and fractional cover approaches

Jan Joseph V. Dida, Arnan B. Araza, Gerald T. Eduarte, Arthur Glenn A. Umali, Pastor L. Malabrigo Jr. & Ramon A. Razal

To cite this article: Jan Joseph V. Dida, Arnan B. Araza, Gerald T. Eduarte, Arthur Glenn A. Umali, Pastor L. Malabrigo Jr. & Ramon A. Razal (2021) Towards nationwide mapping of bamboo resources in the Philippines: testing the pixel-based and fractional cover approaches, International Journal of Remote Sensing, 42:9, 3380-3404, DOI: 10.1080/01431161.2020.1871099

To link to this article: https://doi.org/10.1080/01431161.2020.1871099

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 10 Feb 2021.

Article views: 534

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=tres20
Towards nationwide mapping of bamboo resources in the
Philippines: testing the pixel-based and fractional cover
approaches

Jan Joseph V. Dida\textsuperscript{a,}\textsuperscript{*}, Arnan B. Araza\textsuperscript{b,c,#}, Gerald T. Eduarte\textsuperscript{d},
Arthur Glenn A. Umali\textsuperscript{e}, Pastor L. Malabrig Jr.\textsuperscript{f} and Ramon A. Razal\textsuperscript{f}.

\textsuperscript{a}Institute of Renewable Natural Resources, College of Forestry and Natural Resources, University of the Philippines, Los Baños, Philippines; \textsuperscript{b}Laboratory of Geo-information Science and Remote Sensing, Environmental Sciences Group, Wageningen University and Research, Wageningen, Netherlands; \textsuperscript{c}Environmental Systems Analysis, Environmental Sciences Group, Wageningen University and Research, Wageningen, Netherlands; \textsuperscript{d}Graduate School, University of the Philippines, Los Baños, Philippines; \textsuperscript{e}Department of Forest Biological Sciences, College of Forestry and Natural Resources, University of the Philippines, Los Baños, Philippines; \textsuperscript{f}Department of Forest Products and Paper Science, College of Forestry and Natural Resources, University of the Philippines, Los Baños, Philippines

**ABSTRACT**

In tropical and subtropical countries, the awareness on the importance of bamboos to the environment and economy is increasing and so is the demand for spatial bamboo information. However, mapping bamboos especially those naturally grown has been challenging, as these grasses are often mixed with other land-use and land-cover (LULC). In this study, we used Sentinel 1 and Sentinel 2 remote sensing (RS) images, and their vegetation indices to accurately map the bamboos of Iloilo province in the Philippines using: (1) pixel-based method that mapped bamboos and other LULC at 10 m resolution, and (2) fractional cover method that mapped bamboo cover at 100 m resolution (\% ha\textsuperscript{-1}). A random forest model was trained for each method and then validated per hectare basis using a 50:50 training-validation ratio of a stratified random sample. The fractional cover method showed 0.34 higher Nash-Sutcliffe Efficiency (NSE) and 5.10\% lower Root Mean Square Error (RMSE) than the pixel-based method. Further validation within upland and lowland sites also favoured the fractional cover method, but the results of the two methods were closer in the upland site (bamboo plantation). Errors at 10 m resolution especially in the lowlands were mostly commission errors, likely because of the spectral similarity and proximity between bamboos and > 14 vegetations. Averaging the RS inputs into 100 m resulted in at most 12\% separation of reflectance values among bamboos, forests, and other vegetations. Using the bamboo cover map, a total of 14,795 (+ 1,283) ha bamboos and 7.45 (+ 4.20) million harvestable culms (poles) were estimated for the whole province, where 54\% come from the lowland. We suggest using the fractional cover method for nationwide baselining of bamboo resources.

**CONTACT** Arnan B. Araza \textsuperscript{#} armanara2006@gmail.com Laboratory of Geo-information Science and Remote Sensing, Environmental Sciences Group, Wageningen University and Research, Droevendaalsesteeg 3, 6708 PB, Wageningen, Netherlands

\textsuperscript{*}equal contributors to the study

\textsuperscript{#}Supplemental data for this article can be accessed here.

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
1. Introduction

Bamboos have high ecological and economic value mostly in the tropical and subtropical regions (Lobovikov et al. 2007). Ecologically, bamboos can regulate soil erosion and water flow in catchments, preserve biodiversity, store carbon, and serve as a reforestation species (Ben-zhi et al. 2005; Yiping et al. 2010). Economically, bamboos have multiple uses, either as a raw material or as manufactured products, catering from individual households to industries (Shu & Wang 2015). The growing awareness on the importance of bamboo resources has driven the demand for their spatial information i.e., through bamboo maps (Bystriakova et al. 2003). In countries where naturally grown bamboos are abundant, bamboo mapping is perceived to be challenging because these grasses are often dispersed, mixed with other vegetations and within forest understories (Lobovikov et al. 2007; Wang, Skidmore, and Toxopeus 2009).

Bamboos are often mapped using remote sensing (RS), typically using satellite images. Bamboo mapping using RS is not entirely new as bamboos have been mapped as early as the 1990s. For instance, Saengnin (1993) produced bamboo maps in Thailand by manually interpreting Land Remote-Sensing Satellite System or Landsat images, while Menon (1994) did the same for India, also using Landsat images and their textural properties. While manual image-interpretation was preferred long before, advances in computer technology led to the supervised classification of bamboos (Lobovikov et al. 2007). This software-oriented method involves annotations of land-use and land-cover (LULC) classes such as bamboos, which are matched with the properties of RS images such as multi-spectral bands of Landsat, as the basis for pixel-level mapping. Using such method, the first global bamboo map year 2004 was produced to assess bamboo coverage and biodiversity (Zhao et al. 2017). A decade later, another global map was made to provide global bamboo statistics (Du et al. 2018). These large-scale periodic maps are made possible by the continuity of satellites, i.e., Landsat satellites have been flying for three decades. Longer satellite missions and additional satellites further improved the spatial and temporal (spatio-temporal) resolutions of bamboo mapping. For instance, Zhang et al. (2019) used Landsat time series (30 m) to distinguish the phenology of bamboos, while Li et al. (2019) applied a similar method with the inclusion of Sentinel 2 (10 m) for better phenology analysis, and hence better bamboo classification. Moreover, the use of hyperspectral RS has demonstrated how spectral signatures of bamboos can differentiate them from other LULC classes (Miyasaka et al. 2009; Saini et al. 2014), while the advent of airborne Light Detection and Ranging (LiDAR) has provided structural detail and height of bamboo canopies (Cao et al. 2019).

The advancement in bamboo RS is complemented by machine learning (ML), as ML can be utilized to reveal spatio-temporal patterns between the RS data and bamboo properties, unique enough to classify bamboos among other LULC. For example, Han et al. (2014) used decision trees and nearest neighbour methods to capture the spectral and textural properties of bamboos and classify them at 90% accuracy. They used object-based image classification method that operates on both spectral and geometric features (i.e., bamboo crown) for classification. A similar study by Ghosh and Joshi (2014) used the same properties, this time derived from a 0.5 m resolution WorldView data, and classifiers support vector machine and random forest (RF) to map bamboos with at most 90% accuracy. Aside from object-based classification, they used the pixel-based or ‘hard’
classification that assigns one LULC class per pixel. Aside from spectral and textural properties, Li et al. (2016) added topographic features to train an RF model and accurately predict multi-temporal bamboos at pixel-level. A more complex ML application involved detection of understory bamboos (80% accuracy) using a neural network model based on bamboo seasonality captured from Landsat multi-date images (Linderman et al. 2004). Other ML-aided bamboo studies went beyond classification tasks and focused on biomass/carbon bamboo modelling (Han et al. 2013; Ying et al. 2016). The use of ML for bamboo applications has materialized because of the RS (big) data. This development is timely with the growing interest in assessing bamboo resources for environmental and commercial reasons.

In the Philippines, there is a fledgling bamboo industry that continuously seeks to make a global mark, ranked as the 6th exporter of bamboo products since 2009 (Phimmachanh, Ying, and Beckline 2015; INBAR 2019). In 2010, Executive Order 879 was issued to institutionalize the Philippine bamboo industry and create long-term and research-supported programmes to improve the production, utilization, and marketing of bamboos (Aggangan 2015). Recently, the government has committed to establish 300,000 ha new bamboo plantations by 2022 (ERDB 2019). These economic and environmental programmes all require baselining and monitoring of bamboo resources to be sustainable.

Currently, sources of information about bamboos are mostly non-spatial. These sources include country reports as part of forest resource assessments (FRA); government initiatives such as the Bamboo Information Network; and existing literature with data derived from local and regional forest inventories (Virtucio and Roxas 2003; Lobovikov et al. 2007; Razal 2009; FMB 2018). Existing high-resolution maps in the country could have provided this information gap. However, the official LULC maps in the Philippines are commonly confined with generic land cover classes, while recent LiDAR-based land-use maps only focused on lowland agriculture (Blanco et al. 2015). Global bamboo maps can provide nationwide bamboo numbers (Du et al. 2018) if they are open access and with uncertainty estimates to assess their usability at local scales. The missing spatially explicit information on Philippine bamboos can be generated using state-of-the-art remote sensing.

Mapping the bamboos in the Philippines should consider their spatio-temporal characteristics. First, Philippine bamboos are evergreen, so mapping them based on their temporal properties (i.e., phenology) via RS time series analysis seems inadvisable. Second, most of the Philippine bamboos are clumped (multiple poles/culms in one clump), naturally grown and mixed with other LULC (Rivera 1998). Without bamboo spectral signatures (i.e., from hyperspectral RS), classifying them among neighbouring LULC classes using object-based methods can be challenging (Petropoulos, Vadrevu, and Kalaitzidis 2013). Existing airborne and ground LiDAR data in the Philippines are ideal data sources, but acquiring them is challenged by the transition of government data to open access.

For large-scale heterogeneous landscapes, the use of Sentinel 2 to create vegetation indices and Sentinel 1 to overcome persistent cloud cover while providing backscatter signals of the landscape, can be both utilized towards better LULC mapping (Steinhauser et al. 2018). In terms of the mapping method, the pixel-based has frequently been used for LULC mapping with a bamboo class. An approach that has never been tested at the land-use level with a bamboo class is the fractional cover method. This approach creates
fractional cover maps of a LULC class per unit area (e.g., % open forest ha\(^{-1}\)) based on the proportion of the LULC sub-pixels inside the unit area (Tsendbazar et al. 2018). Although originally implemented at continental to global scales, the fractional cover method also showed potential for LULC mapping in a landscape similar to the Philippines (Seo et al. 2016).

This study aims to accurately map the bamboos and estimate their volume in a bamboo-rich province in the Philippines by testing the pixel-based and fractional cover approaches. Our specific objectives are the following: (1) demonstrate and assess the two methods towards national upscaling of bamboo maps; (2) assess their performance in the upland and lowland bamboo landscapes; and (3) quantify the bamboo area and volume in the province. The relevance of our study can be summarized as a method and data provider for provincial to nationwide baselining and monitoring of bamboo resources.

2. Study area

Our study area is Iloilo province, the largest province (5,001 km\(^{2}\)) in the Western Visayas region, Philippines. The province is situated between 10°20’ to 11°40’ N and 121°50’ to 123°25’ E (Figure 1). The climate within the province falls under the Modified Corona Type I (dry and wet seasons) and Type III or less pronounced dry and wet seasons (Kintanar 1984). The province is dominated by agricultural lands (46%), while the remaining lands

![Figure 1. Map of Iloilo province showing administrative boundaries and neighbouring provinces. The boundary data comes from global administrative areas (GADM) database.](image-url)
are mostly forested mountain ranges; perennial vegetations such as coconut, bananas, and bamboos; and near-coastal lands full of industrial areas, commercial establishments, residential houses, mangroves and fishponds.

**Figure 2** shows the bamboo landscapes in Iloilo either as **Figure 2** (a) ‘upland’ consisting of homogeneous plantations and in patches alongside terraced crops and **Figure 2** (b) ‘lowland’ where bamboos are growing on riverbanks and within built-up areas. Upland areas are mostly public lands while lowlands are those lands that can be privatized.

### 3. Materials and methods

#### 3.1. Overview of the methods

The workflow to map and quantify bamboos in Iloilo is shown in **Figure 3**. First, RS inputs are used to delineate the LULC strata, including a bamboo stratum from a Geographic Information System (GIS) model. Then, random points per stratum are generated and strategically annotated by experts – both LULC classes at 10 m and % bamboo at 100 m. The annotated data plus the bamboo inventory data constitute the reference data, which are equally split for training and validation of the models. Then, random forest (RF)
Figure 3. Workflow diagram of this study highlighting the input-output and analytical processes involved to produce the bamboo maps and density.
classification is used to map bamboos (and other LULC) at 10 m resolution and RF regression for the bamboo cover at 100 m resolution. The maps are compared with the validation data per hectare basis and assessed by four accuracy metrics: coefficient of determination ($R^2$), Nash-Sutcliffe Efficiency (NSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). The same validation procedure is implemented for an upland and lowland site. Lastly, bamboo volume (i.e., total harvestable culms in the province) is computed using Monte Carlo (MC) simulations that used distributions of the provincial bamboo map and the bamboo density data obtained from the field inventory. The steps of the workflow are mostly implemented in R programming interface (The R Core Team 2019).

### 3.2. Remote sensing inputs

Two main RS satellite images were used for this study: Sentinel 1 C-band radar and multispectral Sentinel 2, both commissioned by the European Space Agency. For Sentinel 1, we acquired the dual-polarized Vertical Transmit-Vertical Receive (VV) and Vertical Transmit-Horizontal Receive (VH) Interferometric Wide Swath images at both descending and ascending orbits (Sentinel 1A and 1B). For Sentinel 2, we only selected images with maximum 20% cloud coverage. We obtained preprocessed Level 2 products of Sentinel 1 from January 2018 to December 2019 (59 images) and 20 images of Sentinel 2 from December 2018 (start of Level 2 products) to December 2019 – all from Google Earth Engine or GEE (Gorelick et al. 2017). The pre-processing of Sentinel 1 involved the conversion from Ground Range Detected or GRD (Level 1) to backscatter coefficient unit (decibels, dB) using the standard processing chain of Sentinel 1 as follows: updating the orbit file, removal of GRD and thermal noises, radiometric calibration, and terrain correction. Similarly, the pre-processed Sentinel 2 images involved the conversion from top of the atmosphere reflectance to surface reflectance ($\rho$) using the sen2cor atmospheric correction algorithm. Individual bands from Sentinel 1 (VV and VH) and Sentinel 2 (Blue, Green, Red, Red edge, Near-infrared (NIR), Shortwave-infrared (SWIR)) were used. We then generated the mean and median composite of the multi-temporal images of Sentinel 1 and 2, respectively. Then, several vegetation indices were derived: Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Radar Vegetation Index (RVI). Aside from the satellite bands and vegetation indices, a Digital Elevation Model (DEM) from Shuttle Radar Topography Mission (SRTM) (Jarvis et al. 2008) was used and obtained also from GEE. More information about these RS inputs is shown in Table 1.

### 3.3. Sampling the reference data

Stratified random sampling was undertaken to disperse points for reference data annotation. This sampling design was recommended for study areas with a high imbalance in LULC classes (Chen and Wei 2009). Meaningful strata were first delineated by clustering the RS inputs using K-means algorithm. The K-means clustering works by ‘reshaping’ the RS inputs into a planar (2d) coordinate system and then performs a distance-based similarity measure among the reshaped inputs relative to a strata centroid. Six strata
Table 1. List of RS inputs used in the study, comprising of satellite image bands and vegetation indices (with equations) from Sentinel 1 and Sentinel 2. Their units and original resolution are also shown.

<table>
<thead>
<tr>
<th>Satellite platform</th>
<th>RS input</th>
<th>Unit</th>
<th>Original resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentinel 1</td>
<td>VV</td>
<td>Backscatter intensity (dB)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>VH</td>
<td>Index (0 to 1)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>RVI</td>
<td>Index (0 to 1)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Green</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentinel 2</td>
<td>Red edge (Band 5)</td>
<td>Surface reflectance</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Near-infrared (NIR)</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Shortwave-infrared (SWIR) (Band 11)</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>NDVI</td>
<td>Index (0 to 1)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>EVI</td>
<td>Index (0 to 1)</td>
<td>10</td>
</tr>
<tr>
<td>SRTM-DEM</td>
<td>Elevation</td>
<td>Metres Above Sea Level ('m a.s.l.')</td>
<td>30</td>
</tr>
</tbody>
</table>

were optimally generated, corresponding to the following LULC classes based on visual judgement: forest, croplands, built-up areas, mangroves, inland water, and other vegetations.

Given that a bamboo class is technically a land-use class and often mixed with other land-uses, we generated a separate bamboo stratum using multi-criteria spatial analysis. The idea is similar to the work of Fava and Colombo (2017), wherein they created a bamboo stratum to study the reproductive properties of bamboos during fire season. Here we targeted potential bamboo habitats on rolling (slope) areas, riversides, and plantation areas as informed by the local people during the site visit. We buffered rivers by 250 m per side and combined them with areas below 18% slope, both obtained from the DEM data. Percent slope was used not only to capture potential bamboos, but also to align with the official land classification of the Philippines (above 18% as public lands and below 18% as lands that can be privatized). Non-bamboo areas were removed from the river buffers and rolling slopes using the country land-cover dataset in 2010. Finally, we added bamboo plantations from one municipality (Maasin) to complete the bamboo strata.

Per strata, 350 reference data were randomly assigned, perceived to be sufficient relative to the study area size. The labelling of the reference data was done separately by four experts using false colour composites of Sentinel 2 and base maps from Google Earth and Bing Virtual Map. In addition, local bamboo data were integrated to complete the reference data. These are inventory plots and other bamboo spots within the inventory travel route.

Aside from the LULC labels, we obtained the values of the RS inputs from all reference data. Shown in Figure 4 (a) is the strata map itself and the reference data later on divided into training and validation data, see Figure 4 (b). In the supplemental material, the spectral information per LULC class is shown.

### 3.4. Mapping methods

We used a machine learning model for bamboo mapping – RF algorithm – a non-parametric ensemble model of decision trees (Breiman 2001) using the ranger package in R for faster implementation (Wright and Ziegler 2017). The algorithm operates by first
bootstrapping the training data (i.e., half of the reference data), wherein rows of the training data are selected randomly with replacement. From that subset, trees are grown from a limited set of predictors to assure uncorrelated trees. The process to form a tree starts when a root node is split to form the succeeding nodes based on binary decisions until the final node (leaf node) is reached. Once trees are grown, a prediction per tree is made and predictions of all trees are averaged for continuous data (RF regression) or based on majority vote for discrete data (RF classification). RF is perceived to be unique among other similar ML models since the algorithm can provide importance and sensitivity measures to predictors.

3.4.1. **Pixel-based classification**

RF is widely used as a LULC pixel-level classifier in heterogeneous landscapes (Adam et al. 2014) and even land-uses in the Philippine context (Johnson and Iizuka 2016). In LULC with a bamboo class, RF was used for multi-temporal bamboo mapping (Li et al. 2016).

For our first mapping approach, the pixel-based method, we used RF as a LULC classifier (RF classification). Using half of the reference data for model training, we then predicted bamboos and other LULC classes at the original RS input resolution (10 m). For this classification task, we used GEE for efficient implementation.
3.4.2. Fractional bamboo cover

Our second mapping approach is the fractional cover method, which predicted the fractional bamboo cover (% ha\(^{-1}\)) at 100 m resolution (herein called as bamboo cover). First, each point of the reference data was converted into 1 ha reference polygons by expanding the point to square buffers. Each polygon was discretized by 10 × 10 grid cells to guide the bamboo cover labelling. An average of three clumps covers one grid cell based on the inventory. Moreover, the shape of the bamboo crowns is recognizable from the Google Earth base map. Four experts assigned the bamboo cover (0% to 100%) and if their annotations disagree by > 30% standard deviation (SD), we removed that reference data. A total of 140 reference data were discarded.

The fractional cover method used RF regression because 0% to 100% is a continuous variable. The model predicted bamboo cover at 100 m resolution so the RS inputs were resampled using the mean value inside 1 ha pixels. A binary (bamboo and non-bamboo) raster was included as a predictor to the fractional cover method, as suggested in Masiliūnas et al. (2019). The layer was produced from an RF classifier trained from the 100 m RS inputs. This is to deal with skewness towards 0% because bamboo is just one of the seven strata and there were more polygons with no bamboos.

3.5. Validation for the whole province, upland site and lowland site

Predictions from the RF models of pixel-based (RF classification) and fractional cover (RF regression) were both validated using the same validation data (50% of reference data). For comparability, the predictions from pixel-based were scaled per hectare to have fractional cover units. We counted every 10 m bamboo pixels inside each 1 ha reference polygon i.e., 20 pixels = 20% bamboo ha\(^{-1}\).

Validation or how the predicted bamboo cover compares with the reference data was measured by the following goodness-of-fit metrics: \(R^2\) in Equation (1) and NSE in Equation (2); and absolute % bamboo errors: MAE in Equation (3) and RMSE in Equation (4). The components of the equations are symbolized as: \(N\) as the total validation data, \(R_i\) as the reference bamboo cover, \(P_i\) as the predicted bamboo cover, \(\bar{R}\) as the mean % bamboo cover of all \(R_i\) and \(\bar{P}\) as the mean % bamboo cover for all \(P_i\). We plotted the predicted and reference bamboo cover relative to the 1:1 line to visualize the nature of errors either as systematic or random.

\[
R^2 = \frac{\sum_i (R_i - \bar{R})(P_i - \bar{P})^2}{\sum_i (R_i - \bar{R})^2 \sum_i (P_i - \bar{P})^2}
\]

(1)

\[
NSE = 1 - \frac{\sum (R_i - P_i)^2}{\sum (R_i - \bar{R})^2}
\]

(2)

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |R_i - P_i|
\]

(3)
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_i - R_i)^2}

In addition, two bamboo landscapes were validated to assess how the mapping methods differ between the uplands and lowlands (recall Figure 1). For the upland site, we selected a portion of a bamboo plantation in the municipality of Maasin; while for the lowland site, we pre-identified a community with bamboos in backyards and roadides. For each site, we randomly sampled 100 validation data, independent from the provincial-level data. We further examined the predictions visually to assess the bamboo density within and among the sites.

3.6. Uncertainties from the mapping procedure

The mapping process of bamboos involves the reference data (and its sampling), RS inputs, and RF models that are not error-free and may infuse errors to the predictions. To acknowledge the error, estimates of uncertainties i.e., confidence intervals are needed as a standard mapping procedure. In our case, the standard error (SE) of the RF model was derived and used to calculate the 95% confidence intervals (CI_{95%}) of the bamboo map (B). In Equation (5), B was either subtracted or added to 1.96 \times SE (as lower and upper bound of CI_{95%}, respectively). The CI_{95%} also served as the uncertainty range of the bamboo totals.

CI_{95%} = B \mp 1.96 \times SE

3.7. Bamboo inventory, area and volume

The inventory was conducted in October 2018 within 16 pre-identified areas in the province. Thirty 0.09 m^2 plots were inventoried and geo-located following a random sampling design. Each plot was characterized using the following: stand classification, species diversity, number of clumps and number of harvestable culms. For this study, we used the plot density (culms ha^{-1}) and reported the species diversity. As mentioned, the inventory plots and additional geo-tagged bamboos outside the plots were added to the reference data.

The total bamboo area was derived using the provincial bamboo cover map of the method with relatively higher accuracy. The provincial bamboo area was simply the sum of all bamboo cover pixels (% ha^{-1}) divided by 100. The upland and lowland total bamboo areas were also derived based on slope.

The density data (culms ha^{-1}) from the inventory and the provincial bamboo map were used to estimate the total harvestable culms for the whole province using MC approach. The MC simulations used probability distributions of the bamboo area from the CI_{95%} and the density values from the inventory data. Using those distributions, we simulated bamboo area multiplied by the bamboo density 1000 times to come up with the total bamboo culms in the province. We took the mean of the MC simulations and the SD as an uncertainty range for the total number of culms.
3.8. **Influence of RS inputs to the mapping methods**

RF provides a variable importance measure (VIM), a value given to predictors depending on how they would change the model accuracy whenever permuted (noise). We graphed the relative VIM (%) of the RF models and ranked each predictor from the RS inputs.

We also compared the spectral profiles of the LULC classes between the Sentinel 2 images at the original resolution (10 m) used for the pixel-based method and the 100 m resolution used for the fractional cover method to assess the effect of spatial aggregation. We randomly selected 200 pixels for each LULC class and obtained the mean surface reflectance per class.

4. **Results**

4.1. **Provincial and site-specific validation results**

The validation results of the two mapping methods for the whole province are shown in Table 2. Both methods showed reasonable accuracy results i.e., at least 0.50 NSE and $R^2$, but the accuracy metrics were found higher for the fractional cover method, particularly higher $R^2$ and NSE, and lower MAE and RMSE. A larger gap between the two methods was observed in NSE (0.34) and RMSE (5.10%).

The same validation procedure was performed for the upland and lowland bamboo sites (Table 3). In the lowland site, the discrepancy between the accuracy results of the mapping methods was large, most notable is the lower NSE (0.40) of the pixel-based method. In the upland site, the results of the two methods were close, knowing that uplands have a high bamboo cover, and thus relatively higher MAE and RMSE than that of the lowland site. The gap between the NSE of the two methods was reduced to 0.13 and the MAE was 2.74% apart.

Shown in Figure 5 are the comparisons between the reference and predicted bamboo cover for province-wide in Figure 5 (a) – (b), upland site in Figure 5 (c) – (d), and lowland study sites in Figure 5 (e) – (f) to analyse the nature of errors (systematic or random). For the provincial results, we found systematic errors (bias) in the fractional cover method.

### Table 2. Bamboo cover validation results of the two mapping methods at the provincial level.

<table>
<thead>
<tr>
<th>Method</th>
<th>n</th>
<th>$R^2$</th>
<th>NSE</th>
<th>MAE (% ha$^{-1}$)</th>
<th>RMSE (% ha$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fractional cover</td>
<td>1157</td>
<td>0.85</td>
<td>0.84</td>
<td>3.51</td>
<td>6.54</td>
</tr>
<tr>
<td>Pixel-based</td>
<td>1157</td>
<td>0.56</td>
<td>0.50</td>
<td>6.11</td>
<td>11.64</td>
</tr>
</tbody>
</table>

### Table 3. Validation results of the two bamboo cover mapping methods for the upland and lowland sites.

<table>
<thead>
<tr>
<th>Site, method</th>
<th>n</th>
<th>$R^2$</th>
<th>NSE</th>
<th>MAE (% ha$^{-1}$)</th>
<th>RMSE (% ha$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upland, Fractional cover</td>
<td>100</td>
<td>0.67</td>
<td>0.60</td>
<td>10.16</td>
<td>13.91</td>
</tr>
<tr>
<td>Upland, Pixel-based</td>
<td>100</td>
<td>0.59</td>
<td>0.47</td>
<td>12.90</td>
<td>19.11</td>
</tr>
<tr>
<td>Lowland, Fractional cover</td>
<td>100</td>
<td>0.75</td>
<td>0.70</td>
<td>3.11</td>
<td>4.86</td>
</tr>
<tr>
<td>Lowland, Pixel-based</td>
<td>100</td>
<td>0.57</td>
<td>0.30</td>
<td>3.74</td>
<td>7.67</td>
</tr>
</tbody>
</table>
Figure 5. Results of the comparison between the predicted and reference bamboo cover (% ha$^{-1}$) for (a) province site, fractional cover method; (b) province site, pixel-based method; (c) upland site, fractional cover method; (d) upland site, pixel-based method; (e) lowland site, fractional cover method; and (f) lowland site, pixel-based method. Blue lines correspond to regression lines.
results with overestimation at 25% to 45% bamboo ha\(^{-1}\), while we observed more spread (random errors) and overestimation at 0% and > 70% bamboo ha\(^{-1}\) in the pixel-based results. In the upland site, both methods overestimated the 20% to 50% bamboo ha\(^{-1}\) range especially the pixel-based method. In the lowland site, a better fit was observed in the results of the fractional cover approach.

### 4.2. Provincial bamboo cover map

The fractional cover method was used to derive the provincial bamboo map with 95% confidence interval (Figure 6). From this map, the total area of bamboos in Iloilo was estimated at 14,795 (± 1,283) ha. In Figure 6 (a), it can be observed that bamboos were spread within the whole province and the densest cover were found in the mid-latitudes of the province (corresponding to the location of the municipalities of Maasin, Alimodian, and Janiuay, see Figure 1). Moreover, patches of dense bamboos were found in the southwest bound of the province, while lesser bamboos were observed going northeast. The same province-wide map from the pixel-based method is shown in the supplemental material, where we found a similar spatial pattern of bamboos, but with relatively higher bamboo predictions and spread.

The CI\(_{95}\) showed a narrow range as indicated by (almost) similar bamboo bin colours, see Figure 6 (b) – (c).

![Figure 6](image-url)
The bamboo area in the uplands and lowlands were estimated to be 6,871 ha and 7,924 ha, respectively. Shown in Figure 7 are pixels with bamboo cover categorized as either lowland or upland pixels Figure 7 (a). The figure also shows upland Figure 7 (b) and lowland Figure 7 (c) representative sites with bamboo cover at around 1 km scale. In the upland sample site, we observed variable bamboo cover with patches from low to dense cover. On the other hand, we observed consistent low bamboo cover in the sample lowland area.

4.3. Bamboo volume and species

Using MC simulations of the bamboo density data (culms ha⁻¹) and the provincial bamboo map from the fractional cover method, we estimated 7.45 (± 4.20) million harvestable culms in the province.

Three species belonging to three genera were recorded from the 16 local areas visited: Bambusa spinosa Roxb., Dendrocalamus asper (Schultes f.) Backer ex Heyne, and Gigantochloa levis (Blanco) Merr. The most abundant bamboo species was B. spinosa, covering almost 98% of the total clumps recorded.

Figure 7. Pixels (100 m) with > 0% bamboo cover categorized as either upland or lowland (a); a sample site from an upland bamboo plantation with variable bamboo cover (b); and a community in the lowland that is low in bamboo cover (c).
4.4. Influence of RS inputs to the bamboo cover prediction

The importance of each predictor was determined based on how they affect model accuracy once permutated, also known as the VIM values of RF models. Using the RF regression model, the rank of the relative VIM values is shown in Figure 8. Predictors VV, NDVI, and VH constituted 63% of the total VIM, while the rest (37% of total) were all below 10%, the lowest happened to be DEM, SWIR, and Green predictors.

The comparison of mean surface reflectance between the original and aggregated RS inputs (Sentinel 2) is shown in Figure 9. Compared to the original resolution in Figure 9 (a), the reflectance values of the 100 m resolution in Figure 9 (b) showed more spectral distinction among bamboos, forests, and other vegetation, especially from Red, NIR, and SWIR (bands 4, 8 and 11). Bamboos had lower reflectance (6% and 8%) in the Red band, 4% lower and 1% higher in the NIR band, and 12% higher and 11% lower in the SWIR band than forests and other vegetation, respectively.
5. Discussion

5.1. Validation of the mapping methods

The pixel-based and fractional cover methods were reasonably accurate in mapping bamboos, the latter having a better comparison with the reference data. Both methods exhibited systematic errors (bias), while random errors were more evident from the pixel-based method. The fractional cover method used a regression method (RF regression), which can be prone to the ‘bias to the mean’ phenomenon that exhibits overprediction of the lower values and underprediction of the higher values (Zhang and Lu 2012). On the other hand, the errors from the pixel-based are driven more by the error of commission as more non-bamboo 10 m pixels were classified as bamboos. The spectral profiles of bamboos, forests, crops, and other vegetation overlapped in most bands (especially NIR band), see supplementary material. Specifically, the LULC class ‘other vegetations’ which can be tree orchards, coconuts, brushes, and other woody grasses can be falsely detected as bamboos mainly because of their spectral similarity and proximity to bamboos (Zhang et al. 2019). Similarly, forests misclassified as bamboo pixels may happen when the latter exhibit forest-like signals, particularly mature bamboos with denser and greener culms (de Carvalho et al. 2013). Moreover, crops classified as bamboo pixels may occur when bamboo clumps are not dense enough to avoid mixed signals with soil i.e., due to harvest and age. Misclassifications of bamboos to other LULC can be overcome when either using very high-resolution (VHR) multispectral images such as 0.5 m to 2 m resolution (Ghosh

Figure 9. Spectral profiles of LULC classes from 10 m resolution (a) and 100 m resolution (b) Sentinel 2 bands using randomly selected reference pixels. Notice the overlap between bamboos and other vegetations in (a).
and Joshi 2014; Yu et al. 2020) or hyperspectral RS images to dissect spectral signatures even among hundreds of tropical vegetations including bamboos (Saini et al. 2014). Height data from LiDAR can provide structural classification of bamboos (Cao et al. 2019). However, these datasets are mostly not open source.

In the upland site, both mapping methods showed relatively similar validation results given the high bamboo cover in the landscape. Comparatively, bamboo plantations are more homogeneous and sometimes mixed with terraced crops, wherein in the lowlands, bamboos are mixed with at least 14 agricultural land-uses within the backyards of houses and parallel to shorelines (PPDO 2014).

### 5.2. Mapping method novelty

The estimated bamboo area for the whole province was 14,795 ha – almost twice the reported estimate of Western Visayas Agriculture and Resources Research and Development Consortium (Malaya 2017). The difference between the two estimates implies different methodologies used. Here we used remote sensing images and products, which provide spectral, backscatter intensity, and height characteristics that can distinguish vegetations from one another (Sohl and Sleeter 2011). Fusing Sentinel 1 and Sentinel 2 for LULC mapping can yield more accurate LULC maps instead of just using either one of the satellites (Steinhausen et al. 2018).

With ML, specific patterns from the RS inputs and bamboos can be ‘learned’ and can be extrapolated to the whole province. This sensitivity is also reflected by the VIM values, where both values of radar and optical predictors are high. Sentinel 1 bands showed higher sensitivity constituting 43% of the total VIM. The VV band (highest VIM) likely intercepted areas with 0% bamboo ha\(^{-1}\) such as inland waters, mangroves, and even croplands – classes with relatively low backscatter intensity because of moisture (Wagner et al. 2012; Steinhausen et al. 2018). The bamboo cover reference data are skewed to the left, containing mostly no bamboo cover samples. In landscapes where VV and VH signals are saturated (no sensitivity) due to canopy cover, the optical predictors i.e., NDVI (20% of total VIM) can provide greenness distinction among bamboos, forests, and other vegetations (Zhang et al. 2019).

### 5.3. Spatio-temporal aggregation effect

The reflectance values among bamboos, forests, and other vegetations showed at most 12% difference in the Red, NIR, and SWIR bands when aggregated from (at least) 10 m to 100 m spatial resolution. This aggregation effect reflects again the heterogeneity in the landscape of the study area and justifies why the fractional cover method (100 m) showed higher accuracy results than the pixel-based (10 m) at large-scale mapping. In the context of fractional cover mapping, signals due to mixed pixels are avoided as pixels capture the actual fractions (sub-pixels) of any LULC class inside (Masiliūnas et al. 2019; Latifovic and Olthof 2004). Moreover, hard classification of bamboos at 100 m may lead to overestimation of the provincial bamboos as we seldomly encountered reference data full of bamboos. While spatial aggregation until 6 km was found to provide accurate fractional covers of vegetations (Colin et al. 2018), we believe that the aggregation for bamboo cover mapping should not be way coarser than 100 m (i.e., 1 km onwards). This caveat is
emphasized to avoid losing spatial detail related again to the heterogeneity in the landscape where Philippine bamboos exist. Recall that even bamboo plantations in Iloilo in Figure 7 (b) had pixels with low density because of bare soil and terraced crops at a 1 km scale. We have not found any investigation so far on the effects of different spatial resolutions when mapping fractions of any minority LULC class such as bamboos.

The spatial aggregation is complemented by the temporal aggregation (mean and median composites) of the Sentinels 1 and 2 multi-date images, assuring minimal cloud presence and more credible LULC signals (Carrasco et al. 2019). This favours perennial vegetations such as the bamboo species in the Philippines, which are aerially detectable year-round.

5.4. Uncertainties from the bamboo mapping procedure

Error or the deviation of the predicted to the true bamboo cover is often unknown so an uncertainty (error range) is estimated. Here we acknowledge three prominent error sources, and we emphasize present and future remedies to minimize them. First, while we used a stratified random sample to generate the reference data and filtered them whenever four experts disagree by > 30% SD, the reference data can still be prone to human bias. The annotation process can further be improved by addressing the subjectivity in annotations i.e., through a feedback mechanism among experts (Tarko et al. 2020) and increasing the number of experts (McRoberts et al. 2018). Second, as previously mentioned, errors from the RS inputs are likely minimized after temporal aggregation of Sentinel 1 and Sentinel 2 multi-temporal images. A cloud-free Sentinel 2 composite for the whole province needed image collections of a year, which is reasonable given that bamboos are perennial vegetations and removal of entire clumps are seldom due to the economic value of bamboos. Third, the RF model predictions are essentially the mean of the predictions from all the random trees and thus need sampling considerations. Considering that, the RF implementation we used (R ranger package) automatically provides a standard error of RF models (Wright and Ziegler 2017). Once the bias from RF regression becomes non-negligible come country-wide mapping, we would explore bias removal methods.

5.5. Current and potential bamboo habitats

Dense bamboo plantations in Iloilo were concentrated in the upland municipalities of Maasin, Janiuay, and Alimodian. But the tally of bamboos in the lowlands was slightly higher than that of the uplands, which is directly related to the terrain of the province being 76% lowlands or potential private lands (FMB 2018). Recall also that the upland parcel in Figure 7 (b) depicted low to medium density bamboo cover, which can be indications of active harvesting due to high economic demand. Non-stop harvest of bamboo culms can also be hinted by the high SD of the total harvestable culms. Presently, the proportion of upland-lowland bamboos and the signs of active harvesting are key findings for the stakeholders as far as land management is concerned.

For the on-going and future expansion of bamboos in the uplands (> 18% slope), it would be ideal to plant bamboos in watersheds and along riparian zones (riversides). Recall that we found decreasing bamboos going northeast of the province, areas outside
the watersheds and areas that fall under a different climate type (Kintanar 1984). Moreover, there are multiple-use zones of public lands suitable for co-management between the government and local organizations – a strategy not only for conservation, but also for livelihood purposes (Siva 2004). Using our provincial bamboo map as the main input, a bamboo plantation suitability map can be modelled similar to the GIS-based approach we used to create the bamboo stratum for sampling.

We did not account for bamboos in forest understories because they are relatively few, hardly visible to the satellites, and there are limited inventory data to validate them. But once more forest inventory or ground LiDAR becomes available, bamboo understory fractions can also be mapped.

5.6. Future directions of bamboo resources mapping

Both mapping methods have the potential for upscaling, the pixel-based method though has more room for improvement in mapping lowland bamboos mixed with other land-uses – the majority of Philippine bamboos. Potential strategies to improve the said method for large-area mapping would be to integrate structural and height data coming from airborne LiDAR, and use of the recent agricultural LULC maps derived from VHR images as additional LULC annotation. We could have integrated such existing datasets if the data request process in the Philippines was matured. Assuming high-quality training data, it would be easier to use GEE for pixel-level mapping. We provided a GEE script workflow for that purpose (see code availability).

Once bamboo plantations expand, we foresee that the pixel-based method can be used for monitoring their development assuming they are homogeneously planted. Plantations are more management-oriented and will benefit from high-resolution maps at 10 m (Van Passel, Keersmaecker and Somers 2020). Using these sites and the ‘bamboosetums’ (bamboo parks and research plantations) in the country, several follow-up studies may include species-level mapping and development of bamboo indices (Goswami, Tajo, and Sarma 2010). Similarly, other agro-industrial plantations such as oil palm, rubber tree, and even coconut can be mapped (site-specific) at the original resolution of Sentinel 1 and Sentinel 2. Whenever VHR images (< 5 m) are purchased or become open access, automated counting of trees in plantations is promising (Nomura and Mitchard 2018).

For nationwide bamboo mapping (baselining), we currently suggest the use of the fractional cover method using the free Sentinel 1 and Sentinel 2 RS images. The annotation of bamboo cover is relatively easy since bamboo crowns are recognizable from base maps such as Google Earth and one can save time by not labelling other LULC classes. The annotation of fractional units can be more efficient when using online platforms e.g., GeoWiki (Tsendbazar et al. 2018). Moreover, the advantages of using GEE for data access and preprocessing together with the statistical functions of R can be both utilized using the new rgee package (Aybar et al. 2020). The fractional cover approach can also be applied for baseline and large-scale mapping of other forest trees with economic potential such as Almaciga (Agathis philippinensis Warb., Monsunia), Benguet pine (Pinus kesiya Royle ex Gordon var. langbianensis (A.Chev.) Gaussen ex Bui), Mindoro pine (Pinus merkusii Jungh. & de Vriese), etc.
In terms of nationwide bamboo density, additional inventories are needed to capture the variability of bamboos in terms of stand, species, age, and other key bamboo information. Nationwide inventory may follow a stratified systematic sampling, where participatory information from local entities may help pre-identify bamboo habitats.

6. Conclusion

We demonstrated and compared two bamboo mapping methods using RS inputs and RF models towards nationwide bamboo mapping. While both methods were accurate enough statistically, we suggest using the fractional cover approach whenever large-scale bamboo mapping is aimed i.e., country baselining. This is to compensate for three potential issues in mapping Philippine bamboos: (1) the spectral similarity and proximity of bamboos with other land-uses; (2) difficulty to access hyperspectral and LiDAR data, which could respectively address the previous point; and (3) complexity in annotating LULC classes at the country-level. Aside from baselining, there will be bamboo monitoring activities in line with the 300,000 ha bamboo expansion target of the government. In this case, the spatial detail at 10 m is advantageous so it would be logical to use the pixel-based method particularly if the bamboo plantations are homogeneous (i.e., not inter-cropped).

Slightly higher bamboo total in the lowlands than in the uplands is reasonable as the former consist 76% of the total provincial area. Interestingly, upland bamboo plantations were not entirely high in bamboo cover, which can be indications of active bamboo harvesting. This status of bamboo utilization is also indicated by the relatively high SD of the total harvestable culms. With these evidences, the bamboo expansion plan of the government should materialize to meet current and future demands on bamboo resources.

The provincial bamboo map and density estimates we provide can be used by environmental and business entities in Iloilo. Lastly, the automated steps for both pixel-based and fractional cover methods are open source, and hence reproducible in study areas outside the Philippines.

Acknowledgements

The authors would like to thank PCAARRD-DOST for funding the project. Gratitude is given to the Iloilo foresters and local people who helped the team during the participatory workshops and field inventory. Appreciation is also given to Dainius Masiliūnas of Wageningen University and Research for his technical suggestions on the methodology.

Disclosure statement

No foreseen conflict of interest is reported by all authors.
Funding

This research is funded by the Department of Science and Technology Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (DOST PCAARRD);

ORCID

Jan Joseph V. Dida [1] http://orcid.org/0000-0001-9023-9346
Gerald T. Eduarte [1] http://orcid.org/0000-0002-8147-0981
Arthur Glenn A. Umali [1] http://orcid.org/0000-0001-9289-6649
Ramon A. Razal [1] http://orcid.org/0000-0002-8850-2130

Code availability

https://github.com/arnanaraza/BambooPH

Author contribution

Dida – Conceptualization, Data curation, Funding acquisition, Investigation, Writing inputs, Visualization
Araza – Conceptualization, Data curation, Formal analysis, Methodology, Manuscript writing, Visualization
Eduarte – Data curation, Formal analysis, Writing inputs, Investigation, Formatting, Field work
Umali – Data curation, Field work
Malabrigo – Data curation, Funding acquisition
Razal – Conceptualization, Administration, Supervision, Funding acquisition, Draft review and editing

References


Saini, M., B. Christian, N. Joshi, D. Vyas, P. Marpu, and N. Krishnayya. 2014. “Hyperspectral Data Dimensionality Reduction and the Impact of Multi-Seasonal Hyperion EO-1 Imagery on...


