Logging leaves a fingerprint on the number, size, spatial configuration and geometry of tropical forest canopy gaps

Beibei Zhang | Fabian Jörg Fischer | David A. Coomes | Tommaso Jucker

Abstract
Recent advances in remote sensing such as airborne laser scanning have revolutionized our ability to accurately map forest canopy gaps, with huge implications for tracking forest dynamics at scale. However, few studies have explored how canopy gaps vary among forests at different successional stages following disturbances, such as those caused by logging. Moreover, most studies have focused exclusively on the size distribution of gaps, ignoring other key features such as their spatial distribution and shape. Here, we test a series of hypotheses about how the number, size, spatial configuration, and geometry of gaps vary across a logging disturbance gradient in Malaysian Borneo. As predicted, we found that recently logged forests had much higher gap fraction compared to old-growth forests, a result of having both a greater total number of gaps and a higher proportion of large gaps. Regrowing forests, on the other hand, fell at the opposite end of the spectrum, being characterized by both fewer and smaller gaps compared to nearby old-growth forests. Across all successional stages gaps were found to be spatially clustered. However, logging significantly diluted the degree of spatial aggregation and led to the formation of gaps with much more complex geometries. Our results showcase how logging and subsequent regrowth substantially alter not just the number and size of gaps in a forest, but also their spatial arrangement and shape. Linking these emergent patterns to their underlying processes is key to better understanding the impacts of human disturbance on the structure and function of tropical forests.

Keywords
airborne laser scanning, canopy height models, canopy structure, forest degradation, forest dynamics, remote sensing, selective logging
1 | INTRODUCTION

Gaps are a ubiquitous feature of forest canopies, the mark left behind by one or more trees dying and falling over (Jucker, 2022; Schliemann & Bockheim, 2011). In tropical forests, these canopy openings play a central role in driving forest dynamics (Brokaw, 1985; Dalagnol et al., 2021; Wright et al., 2003), as they allow sunlight to reach the forest floor—often for the first time in decades—providing seedlings in the understory with an opportunity to grow into the canopy (Wright et al., 2003). Canopy gaps are also hubs of biological activity, with large volumes of deadwood attracting decomposers such as termites, which in turn speed up rates of wood decomposition and nutrient cycling (Griffiths et al., 2021). Characterizing the size and spatial distribution of canopy gaps can therefore help us better understand and model how forest structure and dynamics vary across tropical landscapes (Brokaw, 1985; Cushman et al., 2009). Crucially, canopy gaps also provide an opportunity to understand how tropical forests are impacted by human disturbances such as logging and habitat fragmentation, which are pervasive across the tropics and leave a lasting fingerprint on the structure, dynamics, and composition of these ecosystems (Houghton, 2013; Milodowski et al., 2021; Ordway & Asner, 2020; Riutta et al., 2018; Swinfield et al., 2020). However, it is only with the relatively recent advent of remote sensing technologies such as airborne laser scanning (ALS or LiDAR) that it has become possible to accurately measure and map forest canopy gaps at landscape scales (Goodbody et al., 2020; Jucker, Bongalov, et al., 2018; Kellner & Asner, 2009). Consequently, we continue to miss a complete picture of how logging and subsequent forest regeneration impact different features of canopy gaps, such as their number, size, spatial distribution, and shape.

Compared with conventional ground-based surveys, ALS has revolutionized the way we study canopy gaps by allowing researchers to generate high-resolution 3D models of the forest canopy as seen from above (Jucker, 2022; Silva et al., 2019). From these canopy height models, we can easily detect gaps at any height level aboveground and measure their size, number, location, and shape (Silva et al., 2019). This approach has been used by previous studies to show that, as one would expect, logging results in bigger and more numerous gaps (Kent et al., 2015; Wedeux & Coomes, 2015). This shift in the number and size of gaps can also be seen when looking at the gap size-frequency distribution (GSFD) of a forest. GSFDs tend to follow a power law distribution, the scaling exponent of which ($\lambda$) reflects the ratio of small-to-large gaps (Kellner & Asner, 2009). Human-modified forests impacted by selective logging and fragmentation have been observed to have lower values of $\lambda$ compared to their old-growth counterparts, indicating a greater relative frequency of large gaps (Kent et al., 2015; Reis et al., 2022; Wedeux & Coomes, 2015).

Although GSFDs have proven a useful tool for capturing the effects of human disturbances on tropical forest structure, focusing exclusively on the number and size of gaps overlooks other important attributes of canopy openings that are likely to be affected by logging and other disturbances. This includes features such as the spatial distribution of canopy gaps and the complexity of their shape—both of which can be easily captured by ALS (Jucker, 2022). Recent work using spatial point pattern analysis suggests that canopy gaps may tend to be more spatially clustered than we might expect by chance (Silva et al., 2019), which could play an important role in promoting the spatial structuring of tree diversity across tropical landscapes. This tendency for gaps to be spatially aggregated could be due to a number of reasons, such as spatial variation in topography and soil nutrients that shape the rate at which new gaps are formed and closed (Cushman et al., 2022; Goulamoussène et al., 2017; Jucker, Bongalov, et al., 2018), as well as the intrinsic nature of the disturbance process itself. For instance, in forests where the dominant driver of disturbance are localized windthrows caused by microbursts, we might well expect resulting canopy gaps to be spatially clustered around areas hit by previous storms. However, we simply do not know if and how logging impacts the spatial configuration of gaps in tropical forests. On the one hand, by simply increasing the total number of gaps, logging may dilute the degree of spatial aggregation we observe. On the other hand, if logging is highly selective (e.g., by targeting specific tree species that are themselves nonrandomly distributed across the landscape), it could actually increase the degree of spatial aggregation of canopy gaps.

Logging not only generates new gaps across the landscape, but is also likely to change their shape and geometric complexity by altering the way trees die. While some trees die standing due to drought stress, water logging, disease, and resources competition (Chao et al., 2009), others fall over when they die due to destructive disturbances like blow downs and logging, severely damaging and often bringing down neighboring trees in their path (Esquivel-Muelbert et al., 2020; Reis et al., 2022). These differences in mode of mortality should be reflected in the shape and geometric complexity of the canopy gaps that are formed, which in turn are likely to influence microclimatic environments within gaps that directly impact seedling establishment and growth (e.g., light, temperature, wind flow). For instance, we might expect that gaps created by large emergent trees falling over would have a more complex perimeter due to the impacts on surrounding trees, and therefore a higher perimeter-to-area ratio compared to those created from standing dead trees. However, while some work has been done on characterizing gap shapes from ALS using fractal geometry (Bonnet et al., 2015; Garbarino et al., 2012; Vepakomma et al., 2008), to the best of our knowledge no one has applied these approaches to test how gaps in logged forests differ in their geometric properties from those of primary forests.

Finally, most of our knowledge to date on how logging impacts gap patterns is based on simple comparisons between logged and unlogged forests. However, a crucial feature of canopy gaps is their transience. Once a tree has fallen, the resulting open space is rapidly re-colonized by plants growing vertically from below or extending their branches horizontally into the gap. Gap closure rates in tropical forests can reach 30% yr$^{-1}$ and nearly full closure can be achieved within as little as 2–5 years (Asner et al., 2004; Dalagnol et al., 2019).
We would thus expect the number and size of gaps to decrease with time after logging, before slowly approaching predisturbance levels once trees grow tall enough to generate large gaps when they die. Regrowth should also be a more gradual, symmetric, and less random process than disturbance, and thus lead to a more homogeneous canopy, a reduction in the geometric complexity of the gaps—thereby progressively erasing the fingerprint of logging on gap size structure and spatial arrangement. However, secondary forest succession in the tropics is a complex process with no uniform trajectory (Norden et al., 2015), and it is possible that the canopy structure and gap patterns of regrowing forests are distinct from those seen in both recently logged and old-growth forests.

To determine how logging impacts the number, size, spatial configuration, and shape of canopy gaps, here we used ALS data acquired over a landscape in Malaysian Borneo which includes a mix of recently logged forests, regrowing forests that were previously logged and unlogged old-growth forests (Figure 1). Using the ALS data, we first delineated all canopy gaps extending to multiple height levels aboveground (2-10 m) in 18 circular 1 km² plots (i.e., six located in recently logged forests, six in regrowing forests and six old-growth forests), a scale large enough to robustly characterize GSFDs (Lobo & Dalling, 2014). We then used these data to test the following predictions:

1. The total number of gaps and the cumulative gap fraction will increase substantially immediately after logging, but both will decrease again rapidly during regrowth.
2. GSFDs follow a power-law distribution regardless of logging history, but logging will result in a lower scaling exponent (λ), indicating a greater relative frequency of large canopy gaps. λ will increase again during regrowth, as larger logging-induced gaps are rapidly closed and replaced by smaller gaps caused by natural tree deaths.
3. Canopy gaps in old-growth forests tend to be spatially clustered, especially larger ones reaching all the way down to the forest floor. However, logging will weaken or even erase this spatial pattern, with limited recovery of spatial clustering during the initial regrowth phase due to an absence of large trees that have the potential to cause large gaps.
4. Logging will increase the geometric complexity of gaps by creating canopy openings with high perimeter-to-area ratios. Regrowth will subsequently reduce this complexity as large logging-induced canopy gaps are progressively filled from below and the sides.

2 | METHODS

2.1 | Study sites and sampling design

The study was conducted in Sabah, Malaysian Borneo (Figure 1). The region’s climate is tropical, with a mean annual rainfall of 2700mm and a mean annual temperature of 26.7 °C (Kumagai & Porporato, 2012; Walsh & Newbery, 1999). Sabah’s lowlands were originally dominated by tall and structurally complex dipterocarp forests (Bryan et al., 2013; Jucker, Bongalov, et al., 2018). The dominant drivers of natural disturbance in these forests are localized windthrows caused by downbursts (strong descending winds associated with severe convective systems) and landslides on steep terrain (Jucker, Bongalov, et al., 2018; Riutta et al., 2018). Additionally, periods of drought associated with El Niño events are also known to cause tree mortality, particularly in forests growing on ridge tops where soils are shallower (Miyamoto et al., 2021; Nunes et al., 2021) and near forest edges (Qie et al., 2017). Starting in the early 1970s large swaths of Sabah were logged and converted to rubber and oil palm plantations, so that today only 59% of the region is covered by the majority of which have been selectively logged (Asner et al., 2018; Gaveau et al., 2014; Riutta et al., 2018). Commercial logging practices in the region primarily target mature trees (>60 cm in stem diameter) belonging to the Dipterocarpaceae family, which are conventionally harvested using heavy machinery and then extracted through skid trails (Bryan et al., 2013; Pinard & Putz, 1996).

In 2011, the Stability of Altered Forest Ecosystems (SAFE) project was established to better understand the impacts of logging and forest fragmentation on biodiversity and ecosystem functioning (Ewers et al., 2011). Within the SAFE landscape, tropical forests have been logged multiple times and left to recover at different time points. These logged forests are contrasted with old-growth forests that have never been logged at nearby Maliau Basin and Danum Valley conservation areas, which serve as baseline against which to test the impacts of logging on forest structure, function, and biodiversity (Ewers et al., 2011).

For our study, we subdivided the SAFE landscape into three broadly defined forest types, each of which is represented by six circular 100 ha plots (see Figure 1 and Table 1 for details): (a) Recently logged forests in the core SAFE landscape that have undergone between 2 and 4 rounds of logging up until 2008 (orange circles in Figure 1); (b) Regrowing forests that have been selectively logged twice up until the 1980–1990s and then left to recover (blue circles in Figure 1); and (c) Old-growth forests that have never been previously logged before at Maliau and Danum (three plots at each site; gray circles in Figure 1). Overall, this results in a total of 18 × 100 ha circular plots. Recently logged and regrowing forests were subjected to similar logging intensities in the first round, but the subsequent rounds of logging were lighter in the regrowing forest areas due to restrictions on timber quotas (Struweberg et al., 2013). Moreover, forests surrounding these recently logged plots have either been converted to oil palm or are currently in the process of being converted. Plot locations in the recently logged forests match those of the SAFE project experimental design, while plots in the regrowing and old-growth forests were placed at random within the areas surveyed by ALS (described below), excluding major rivers and roads. Besides differing in their logging history, the study plots share similar soil properties (Riutta et al., 2018) and occupy a relatively narrow elevational range band (approximately 300–600 m a.s.l.; see Table S1 in Supporting Information for details of topographic variation within and among plots).
2.2 Airborne laser scanning (ALS) data acquisition and processing

In 2014, ALS data were collected using a Leica LiDAR50-II sensor across the SAFE project landscape, as well as at Maliau Basin and Danum Valley (see Jucker, Asner, et al., 2018 for details on data acquisition and processing). The software LAStools was used to first normalize the ALS point clouds and then generate 1-m resolution canopy height models (CHMs) in raster format for each site, where each grid cell of the raster layer corresponds to the height of the canopy.
Table 1: Summary attributes of the 18 study plots. Canopy heights (mean ± 1SD) derived from the airborne laser scanning (ALS) data acquired in 2014. Gap fraction, number of gaps km⁻², gap size frequency distribution (GSFD) scaling exponent (λ), Clark-Evans index of spatial aggregation (R), and fractal dimension (D) were measured at the mid-way height level of 6 m aboveground.

<table>
<thead>
<tr>
<th>Plot ID</th>
<th>Location</th>
<th>Forest type</th>
<th>Canopy height (m)</th>
<th>Gap fraction (%)</th>
<th>Number of gaps (km²)</th>
<th>GSFD (λ)</th>
<th>Clark-Evans index (R)</th>
<th>Fractal Dimension (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OG-1</td>
<td>Maliau Basin</td>
<td>Old-growth</td>
<td>47.1 ± 13.8</td>
<td>0.4</td>
<td>93</td>
<td>1.97</td>
<td>0.74</td>
<td>1.33</td>
</tr>
<tr>
<td>OG-2</td>
<td>Maliau Basin</td>
<td>Old-growth</td>
<td>38.5 ± 12.2</td>
<td>0.8</td>
<td>115</td>
<td>1.84</td>
<td>0.72</td>
<td>1.33</td>
</tr>
<tr>
<td>OG-3</td>
<td>Maliau Basin</td>
<td>Old-growth</td>
<td>42.5 ± 14.8</td>
<td>0.7</td>
<td>120</td>
<td>2.74</td>
<td>0.56</td>
<td>1.35</td>
</tr>
<tr>
<td>OG-4</td>
<td>Danum Valley</td>
<td>Old-growth</td>
<td>37.7 ± 17.1</td>
<td>3.8</td>
<td>330</td>
<td>1.80</td>
<td>0.79</td>
<td>1.32</td>
</tr>
<tr>
<td>OG-5</td>
<td>Danum Valley</td>
<td>Old-growth</td>
<td>39.0 ± 16.9</td>
<td>3.3</td>
<td>237</td>
<td>1.79</td>
<td>0.73</td>
<td>1.34</td>
</tr>
<tr>
<td>OG-6</td>
<td>Danum Valley</td>
<td>Old-growth</td>
<td>39.4 ± 14.8</td>
<td>2.1</td>
<td>193</td>
<td>1.78</td>
<td>0.69</td>
<td>1.28</td>
</tr>
<tr>
<td>LF-1</td>
<td>SAFE landscape</td>
<td>Regrowing</td>
<td>28.2 ± 6.9</td>
<td>0.3</td>
<td>46</td>
<td>1.95</td>
<td>0.71</td>
<td>1.32</td>
</tr>
<tr>
<td>LF-2</td>
<td>SAFE landscape</td>
<td>Regrowing</td>
<td>28.8 ± 9.0</td>
<td>2.4</td>
<td>241</td>
<td>1.87</td>
<td>0.76</td>
<td>1.38</td>
</tr>
<tr>
<td>LF-3</td>
<td>SAFE landscape</td>
<td>Regrowing</td>
<td>25.2 ± 7.4</td>
<td>1.3</td>
<td>180</td>
<td>1.75</td>
<td>0.64</td>
<td>1.32</td>
</tr>
<tr>
<td>LF-4</td>
<td>SAFE landscape</td>
<td>Regrowing</td>
<td>27.0 ± 6.8</td>
<td>0.4</td>
<td>70</td>
<td>2.45</td>
<td>0.61</td>
<td>1.30</td>
</tr>
<tr>
<td>LF-5</td>
<td>SAFE landscape</td>
<td>Regrowing</td>
<td>28.6 ± 7.5</td>
<td>0.1</td>
<td>52</td>
<td>2.87</td>
<td>0.79</td>
<td>1.28</td>
</tr>
<tr>
<td>LF-6</td>
<td>SAFE landscape</td>
<td>Regrowing</td>
<td>26.9 ± 6.9</td>
<td>0.4</td>
<td>75</td>
<td>1.89</td>
<td>0.68</td>
<td>1.30</td>
</tr>
<tr>
<td>A</td>
<td>SAFE landscape</td>
<td>Recently logged</td>
<td>14.8 ± 7.6</td>
<td>17.7</td>
<td>384</td>
<td>1.69</td>
<td>0.88</td>
<td>1.36</td>
</tr>
<tr>
<td>B</td>
<td>SAFE landscape</td>
<td>Recently logged</td>
<td>18.7 ± 9.2</td>
<td>8.9</td>
<td>538</td>
<td>2.10</td>
<td>0.94</td>
<td>1.37</td>
</tr>
<tr>
<td>C</td>
<td>SAFE landscape</td>
<td>Recently logged</td>
<td>8.8 ± 7.4</td>
<td>49.9</td>
<td>276</td>
<td>1.59</td>
<td>0.76</td>
<td>1.40</td>
</tr>
<tr>
<td>D</td>
<td>SAFE landscape</td>
<td>Recently logged</td>
<td>11.9 ± 7.9</td>
<td>30.8</td>
<td>495</td>
<td>1.69</td>
<td>0.85</td>
<td>1.40</td>
</tr>
<tr>
<td>E</td>
<td>SAFE landscape</td>
<td>Recently logged</td>
<td>18.2 ± 9.6</td>
<td>12.6</td>
<td>480</td>
<td>1.98</td>
<td>0.88</td>
<td>1.36</td>
</tr>
<tr>
<td>F</td>
<td>SAFE landscape</td>
<td>Recently logged</td>
<td>17.3 ± 9.1</td>
<td>15.4</td>
<td>490</td>
<td>1.90</td>
<td>0.96</td>
<td>1.34</td>
</tr>
</tbody>
</table>

2.3 Canopy gap detection

From the CHMs, we used the ForestGapR package (Silva et al., 2019) to extract the number and sizes of all canopy gaps at multiple height levels aboveground, starting from those extending all the way to 2 m from the forest floor (which is the conventional definition of a canopy gap according to Brokaw, 1982). The approach works by identifying groups of pixels in the CHM that are below a designated threshold height (see Figure 2 for an illustration of the method). Only gaps ≥9 m² were included in the analyses as smaller gaps mostly reflect spaces between tree crowns rather than gaps created by tree fall (Kent et al., 2015). To avoid accidentally missing large gaps, we set the maximum gap size threshold to 100ha, which is the same size as the plots. Based on an initial inspection of the data, we found that in recently logged forests individual gaps started to coalesce into extremely large gaps from around 10 m aboveground due to the almost total absence of mature trees. We therefore restricted further analyses to gaps detected between the height level of 2-10 m aboveground, which we divided into 1-m intervals.

2.4 Canopy gap metrics

2.4.1 Canopy gap number and gap fraction

In addition to counting the number of gaps detected in each plot and at each 1-m height level, we also calculated a measure of cumulative gap fraction. This was done by simply summing the area of all gaps ≥9 m² and then dividing the total by the area of the plot (100ha).

2.4.2 Gap size frequency distributions

To characterize GSFDs, we used the powRlaw package (Gillespie, 2015) to fit a power-law distribution where the frequency of gaps of size x is proportional to: f(x) ∝ x⁻λ (x ≥ xmin , Clauset et al., 2009), where xmin is the lower cut-off value estimated to maximize the fit of the model. The scaling exponent of the power-law function (λ) reflects the ratio of small-to-large gaps, and as a rule of thumb λ < 2 has been suggested an indicator of large, frequent disturbances (Asner et al., 2013; Jucker, 2022).
To ensure we had an adequate number of gaps to estimate $\lambda$ robustly, we restricted this analysis to height level in which we detected a minimum of 20 gaps per $km^2$ (the same threshold was also used for subsequent analyses on gap spatial distributions and geometry). Across the 18 plots, we only identified five instances (3% of the total) in which $<20$ gaps were detected at a particular height level. All five of these occurrences were in plots classified as regrowing and in the height levels closest to the ground (2–3 m aboveground). To determine how well GSFDs are actually fit by power-law functions, we also used the null hypothesis test in the poweRlaw package to compare the goodness of fit of a power-law to that of a log-normal distribution.

2.4.3 | Spatial patterns of forest canopy gaps

We used the Clark-Evans index ($R$) to characterize the spatial arrangement of forest canopy gaps within each 100 ha plot. $R$ is a measure of spatial aggregation of a spatial point pattern, ranging from $R < 1$ when the distribution is clustered or aggregated, $R = 1$ when the distribution is random, and $R > 1$ when the distribution is even or dispersed (Law et al., 2009). To calculate $R$ for each plot and height level, we first used the sf package (Pebesma, 2018) to determine the centroid location of all canopy gaps. We then used the spatstat package (Baddeley et al., 2016) to generate a spatial point pattern from the centroids and then calculate $R$ using the clarkevans function. An edge correction using the cumulative distribution function method was applied to avoid $R$ values being positively biased.

2.4.4 | Canopy gap shape and geometry

To assess whether logging alters the shape of canopy gaps, we used a measure of fractal dimension ($D$) to quantify the complexity of gap geometries for all height levels across the 18 plots (Halley et al., 2004). $D$ was estimated using the perimeter-area method, where larger $D$ values indicated gaps with more complex geometries (e.g., numerous folds; Sugihara & May, 1990). Following the approach of Staver et al. (2019), the relationship between gap area ($A$) and perimeter ($P$) can be expressed as a power-law: $A \propto P^D$, from which fractal dimension can be estimated as $D = \frac{2}{\beta}$. To estimate $\beta$, we used standard major axis regression as implemented in the lmee package (Legendre & Oksanen, 2018) to model the relationship between $A$ and $P$ on a log-log scale for gaps located within each plot (Staver et al., 2019, see Figure S1).
As an alternative measure of gap shape complexity we also tested the isoperimetric quotient (Q), which compares a gap’s observed cross-sectional area (A) to the area of a circle with the same perimeter (P) as follows: \( Q = \frac{4\pi A}{P^2} \) (Nolke et al., 2015). A perfectly circular gap would have \( Q = 1 \), while the smaller the ratio the more irregular the shape of the gap. When we compared average values of Q calculated across all gaps within a plot to values of D, as expected we found that the two metrics of gap complexity were significantly negatively correlated with one another (Pearson’s correlation coefficient = −0.31, \( P < 0.001 \); Figure S2). For this reason we chose to focus on D as the primary metric of gap shape complexity in all subsequent analyses, but present the corresponding results for Q in Supporting Information (Table S2–S3 and Figure S3).

2.5 | Statistical analyses

We used linear mixed-effects models to test how gap number, gap fraction, size-distribution (\( \lambda \)), spatial arrangement (R), and geometry (D) vary between recently logged, regrowing and old-growth forests (factor with three levels), height level (treated as a continuous variable) and their interaction (allowing the relationship between each response variable and height level to vary among the different forest types). To account for repeated measures of gap metrics at multiple height levels in each plot, we included a plot identifier as a random intercept term in the model (factor with 18 levels). Models were fit using the lmerTest package, which makes use of the lme4 package (Bates et al., 2015; Kuznetsova et al., 2017), and to meet the assumptions of normality, all five response variables and height level were log-transformed prior to model fitting. A type III ANOVA was used to determine whether the inclusion of the interaction term between forest type and height level was justified. To summarize the results of each model, we then used post hoc contrasts implemented using the emmeans package (Lenth et al., 2022) to compare the marginal means of the different forest types and calculate Tukey-corrected \( P \)-values for each pair-wise comparison. Similarly, the emtrends function was used to estimate the slope coefficient between each of the five response variables and height level for the three forest types.

3 | RESULTS

3.1 | ALS-derived vertical distributions of canopy gap area fraction and number

In recently logged forests, we found a strong increase in the cumulative gap fraction and total number of gaps compared to old-growth forests (Figure 3a,b and Table 2). Conversely, regrowing forests were more similar to old-growth forests and in fact showed a deviation in the opposite direction, exhibiting a decrease in both cumulative gap fraction and total number of gaps and (Figure 3a,b and Table 2). As expected, we observed an overall increase in gap fraction and total gap number when progressing upward through the height levels (Figure 3a,b), with differences between forest types remaining consistent across height levels (Table 3). However, the rate of increase in gap fraction and number of gaps with height (i.e., the slope of the model) differed between forest types, resulting in a significant interaction term between forest type and height level in the model (see Table S2 for ANOVA output).

3.2 | Impacts of logging on gap size frequency distributions (GSFDs)

Generally, power-law patterns fit GSFDs well for every height level across the three forest types, although log-normal distributions generally provided a similarly good fit to the data (see Tables S4 and S5 for a comparison of the fits the two distributions). Specifically, only in 7% of cases (11 of 157 height levels spanning all three forest types) did we find that a log-normal distribution was a better fit to the data than a power-law, while in other cases there was no clear difference between the two distributions. In terms of differences in GSFDs between forest types, \( \lambda \) values were generally quite similar at lower height levels in the canopy but departed significantly when progressing upward through the canopy (Figure 3c and Table 2). Overall, recently logged forests had lower mean \( \lambda \) estimates compared to old-growth forests (1.81 and 1.93, respectively), indicating a higher proportion of large gaps, and in both cases \( \lambda \) decreased with height level (Table 3). In contrast, mean \( \lambda \) estimates were highest in regrowing forests (2.14) and showed a tendency to increase with height level.

3.3 | Impacts of logging on the spatial arrangement of canopy gaps

Across all forest types and irrespective of the height level at which they were delineated, canopy gaps showed a tendency to be more spatially clustered than expected by chance (i.e., Clark-Evans \( R < 1 \); Figure 3d). However, we found a strong effect of recent logging on the spatial distribution of gaps, which tended to be less spatially clustered compared to old-growth forests (Figure 3d and Table 2). In contrast, gap patterns in regrowing forests were indistinguishable from old-growth forests. The relationship between \( R \) and height levels was consistent across all three forest types, with spatial clustering tending to decrease as one moves up through the canopy (i.e., \( R \) value tending closer to 1; Table 3).

3.4 | Impacts of logging on canopy gap geometry

Fractal dimension (D) was generally much higher in recently logged forests compared to old-growth ones, indicating more complex gap geometries (Figure 3e and Table 2). Regrowing forests were instead statistically indistinguishable from old-growth ones, despite exhibiting a small decrease in D. The height at which gaps were measured did not affect this result qualitatively, but the differences became
more pronounced when comparing gaps delineated higher up in the canopy. Moreover, gap geometric complexity tended to decrease with height levels across all three forest types (Table 3).

4 | DISCUSSION

Our study of tropical rain forests in Malaysian Borneo shows that logging leaves a clear imprint on the three-dimensional structure of the canopy which can be detected when looking at the size distribution, spatial arrangement, and shape of canopy gaps. Logged and degraded forests at different stages of regeneration now cover much of the landscape not only in Borneo but also across the rest of the tropics (Asner et al., 2009; Gaveau et al., 2014; Milodowski et al., 2021; Riutta et al., 2018). Characterizing the canopy structure of these ecosystems and understanding the underlying processes that shape its variation across landscapes is therefore essential to guide their conservation and understand their potential to store carbon (Bongers et al., 2015).

4.1 | Logging creates numerous large gaps in forest canopies

Both gap number and gap fraction increased dramatically after logging, irrespective of the height levels at which they were measured. This was accompanied by a reduction in mean canopy height...
from 41 m in old-growth forests to 15 m in recently logged forests (Figure 1e–g and Table 1), reflecting the almost complete loss of mature trees and a large part of the upper canopy that they dominate (Figure 3a,b, Lobo & Dalling, 2014, Reis et al., 2022). As expected, this pattern was reverted during the first 20–30 years of forest regrowth. During this regrowth phase, mean canopy height increased back to 27 m, and the total number of gaps and the cumulative gap fraction in regrowing forests dropped dramatically, becoming even lower than what we observed in nearby old-growth forests (Figure 3 and Table 2).

This strong decline in canopy gaps, to levels smaller than those observed in old-growth forests, is precisely what we would expect to see in forests undergoing secondary succession following a strong disturbance event (Amaral et al., 2019). Rapid growth into gaps both through lateral branch extensions, infilling from below by seedlings and the likely proliferation of lianas may, for example, result in a more homogeneous vertical structure and a more uniform tree height than expected when comparing to nearby primary forest canopies (Guariguata & Ostertag, 2001). This pattern is also consistent with the high levels (in both absolute and relative terms) of net primary productivity allocated to the growth of new woody components (stems and branches) in logged forests, and effects that persist for decades after the logging activities have ended (Riutta et al., 2018). Another potential explanation for the fewer and smaller gaps observed in regrowing forests is that logging removes primarily the large emergent trees typical of Borneo’s heterogeneous old-growth forests, meaning that once the large canopy gaps created by the removal of these trees have been closed there are no trees left to create new large gaps (Lobo & Dalling, 2013). This is particularly relevant for gap formation, as trees are primarily vulnerable to wind disturbances when exposed above the canopy (Seidl et al., 2014). Furthermore, a greater abundance of lianas in disturbed forests may further reduce the risk of severe wind damage during the initial phases of regrowth (Garrido-Pérez et al., 2008), as vines homogenize the structure of the canopy and help buffer it against strong winds.

### 4.2 Large gaps are prevalent in recently logged forest, but scarce in regrowing forests

When using gaps as an indicator of disturbance, it is common practice to fit a power law function to model GSFDs, as this allows disturbance regimes to be compared using a single metric, the scaling exponent $\lambda$. (Asner et al., 2013; Espírito-Santo et al., 2014; Goodbody et al., 2020; Goulamoussène et al., 2017). Across our study sites we observed a wide spectrum of $\lambda$ values—ranging from 1.68 to 2.74 in old-growth forests, 1.74 to 3.41 in regrowing forests, and 1.56 to 2.10 in recently logged forests—comparable to values reported previously in the literature for both intact and logged forests in the wet tropics (Boyd et al., 2013; Kellner & Asner, 2009; Lobo & Dalling, 2014; Reis et al., 2022). In particular, our results agree with those from a previous study on logged forests in Borneo, which reported $\lambda$ values between 1.66 and 3.76 using a modified power law function (Wedeux & Coomes, 2015).

As expected, the felling of large trees typical of commercial logging operations increased the relative frequency of large gaps, leading to smaller $\lambda$ values in logged forests compared to old-growth ones. However, this effect could only be reliably discerned at the upper height levels at which we delineated gaps (10 m, Figure 3c), which may be due to rapid regrowth following gap formation close to the ground. More generally, regrowth seemed to affect GSFDs as strongly as logging, with even larger $\lambda$ values in regrowing forests than in old-growth ones—a similar pattern as noted for gap fraction and gap numbers. Again, a plausible explanation for this is a combination of an absence of large canopy trees and the rapid turnover of fast-growing species that compete for the newly available space and create small, localized tree fall events in the first decades of succession (Guariguata & Ostertag, 2001).

It is of course also possible that some of the differences we observe between recently logged and regrowing forests could be explained by factors other than the amount of time forests have had to recover. In regrowing forests, for example, the second round of logging was of generally of lower intensity compared to that experienced by recently logged forests, some of which also went through additional rounds of selective logging. These differences in logging intensity and frequency may have accentuated the degree to which the recently logged forests differ from the regrowing and old-growth forests. Additionally, other factors aside from logging can influence variation in canopy structure across tropical landscapes. For instance, topography has been shown to play an important role in shaping the height, gap fraction, and GSFDs of tropical forests in both Sabah and Panama (Jucker, Bongalov, et al., 2018; Lobo &

### TABLE 3 Trends (slope coefficients ± 95% confidence intervals) between each of the five canopy gap metrics and height level across the three forest types.

<table>
<thead>
<tr>
<th>Gap metric</th>
<th>Recently logged</th>
<th>Regrowing</th>
<th>Old-growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap fraction</td>
<td>0.178 (0.159–0.198)</td>
<td>0.257 (0.238–0.276)</td>
<td>0.243 (0.224–0.262)</td>
</tr>
<tr>
<td>Number of gaps (km$^{-2}$)</td>
<td>−0.024 (−0.051–0.004)</td>
<td>0.251 (0.224–0.279)</td>
<td>0.167 (0.139–0.194)</td>
</tr>
<tr>
<td>GSFD ($\lambda$)</td>
<td>−0.013 (−0.021 – −0.006)</td>
<td>0.006 (−0.003–0.015)</td>
<td>−0.005 (−0.013–0.002)</td>
</tr>
<tr>
<td>Clark-Evans index (R)</td>
<td>0.006 (−0.005–0.017)</td>
<td>0.051 (0.038–0.064)</td>
<td>0.051 (0.040–0.061)</td>
</tr>
<tr>
<td>Fractal dimension (D)</td>
<td>−0.003 (−0.006 – −0.001)</td>
<td>−0.004 (−0.007 – −0.001)</td>
<td>−0.009 (−0.011 – −0.006)</td>
</tr>
</tbody>
</table>

Note: A positive slope indicates that a particular metrics tends to increase when progressing vertically up the canopy, while a negative slope indicates a tendency to decrease. Slope coefficients were estimated using the emtrends function applied to the fitted linear mixed-effects models.
Dalling, 2013). This may help explain the differences we see in λ values between the two old-growth sites in our study (Table 1), as the plots at Maliau are generally found on steeper terrain compared to those at Danum (Table S1).

Finally, our results also have more general implications for statistically modeling GSFDs. Although using power law scaling exponents to summarize GSFDs is attractive, it may not always provide the best fit to the data. Fitting power laws is not trivial (White et al., 2008), and the “fat tail” of power law functions can, for example, overestimate the frequency of large gaps (Kent et al., 2015; Wedex & Coomes, 2015). Furthermore, processes underlying power laws are very similar to the generative processes of log-normal distributions, so differentiating between the two is often difficult (Mitzenmacher, 2004). In our study, a log-normal distribution performed similarly well (or even better in some cases) than a power law, fitting a wider range of data. It remains unclear how common this phenomenon might be and a more comprehensive analysis of the distributional characteristics of GSFDs across different forest types would be needed to tackle this knowledge gap (De Lima et al., 2013; Jucker, 2022).

### 4.3 Logging dilutes the degree of spatial clustering of canopy gaps

Natural disturbances of forests often show clustered patterns, both spatially and temporally, that are reflected in the aggregation of canopy gaps (Chambers et al., 2013; Silva et al., 2019). In Borneo, the dominant drivers of natural disturbance are localized windthrows caused by downbursts and landslides on steep terrain (Jucker, Bongalov, et al., 2018; Riutta et al., 2018). Additionally, periods of drought associated with El Niño events can also result in increased tree mortality, particularly in forests growing on ridge tops where soils are shallower (Nunes et al., 2021) and near forest edges (Qie et al., 2017). These processes would generally be expected to lead to aggregated patterns of gap formation such as those we observe at both old-growth sites in our study, as they are both localized in nature and occur nonrandomly across the landscape. By contrast, Borneo lies outside the path of tropical storms, which is one of the reasons why it harbors such tall trees (Jucker, Bongalov, et al., 2018; Shenkin et al., 2019). So unlike other areas of the tropics, it is not subjected to regular severe disturbances that simultaneously affect large tracks of land, which we might instead expect to lead to a more random spatial distribution of canopy gaps.

While logging may induce its own spatial patterns in canopy gaps, we would not necessarily expect these to follow those generated by natural disturbances. Our results suggest this is indeed the case, with more recently logged forests standing clearly apart when comparing the degree of spatial aggregation of their gaps to those of regrowing and old-growth ones. While the degree of aggregation decreased across height levels in old-growth and regrowing forests (i.e., increasing Clark-Evans index R, Figure 3d), R remained close to 1 in logged forests throughout the vertical profile. However, it is difficult to know whether this is a general feature of logged tropical forests or a more specific pattern that characterizes the selective logging operations in Sabah. Commercial logging operation in Borneo tend to preferentially target large, emergent trees belonging to the Dipterocarpaceae family (Riutta et al., 2018). If these trees are themselves not spatially structured across the landscape, then we might expect that logging would lead to a dilution in the spatial aggregation of canopy gaps, exactly like the one we observe in our data. However, selective logging practices vary considerably across the tropics, so it remains to be tested whether logged forests in other regions of the tropics also show this tendency to have gaps that are randomly distributed across the landscape.

### 4.4 The geometry of gaps tends to be more complex in logged forests

We found that recently logged forests had the most complex gap geometries, as reflected by their higher perimeter-to-area ratios. This pattern is likely to reflect the fact that commercial logging operations in Sabah, which use heavy machinery to fell and transport trees using conventional methods, can have a big impact on the canopy by damaging surrounding trees (Matangaran et al., 2019; Roopind et al., 2018; Shenkin et al., 2015). This contrasts with natural tree mortality events, which do not occur exclusively through tree falls (e.g., standing dead trees; Chao et al., 2009, McDowell et al., 2018). Similarly to the other gaps metrics considered in this study, forests recovering from logging seemed to deviate from old-growth patterns in the opposite direction of recently logged forests. Plant growth is a more homogeneous, symmetric, and gradual process than disturbance (Hilty et al., 2021), so this may result in a gradual smoothing of the complex edges created from tree falls.

Fractal geometry has been used extensively in ecology, including to describe spatial and temporal patterns of tree-clustering (Staver et al., 2019) and forest fragmentation (Taubert et al., 2018), but it has not been widely used to characterize gap shape and complexity. In previous studies, gap geometry has been described mostly through idealized geometric forms such as ellipses (De Lima et al., 2013; Garbarino et al., 2012; Goodbody et al., 2020; Vepakomma et al., 2008), but explicitly characterizing the perimeter-to-area ratio may provide new insights on the impacts of logging and other disturbances on ecosystem structure and function. In particular, there are well-known edge effects on tree mortality (Laurance et al., 2011), microclimate shifts (Hofmeister et al., 2019), and biodiversity loss (Püttker et al., 2020; Razafindratsima et al., 2018), which may be increased by logging operations and which directly relate back to gap shape complexity. What remains to be tested is whether different logging practices leave a distinct fingerprint on the shape complexity of resulting gaps. For instance, we might expect that reduced impact logging, which aims to extract timber in ways that minimize damage to neighboring trees (Griscom et al., 2019; Pinard & Putz, 1996; Putz et al., 2008), would result in gaps that have a lower complexity that...
those generated through the more conventional logging practices used at our study sites.

5 | CONCLUSIONS

Our analyses revealed a very clear difference in the frequency, size, shape, and spatial arrangement of gaps between recently logged and old-growth forests. On the surface, regrowing forests instead appeared to be much closer in their gap patterns to old-growth ones, reflecting the rapid regeneration from below and from the surrounding canopy trees. However, for most metrics, the regrowing forest did not only revert to the old-growth pattern, but appeared to deviate in the opposite direction. This suggests that regrowing secondary forests have a qualitatively different structure compared to both recently logged forests and primary forests, with a highly homogeneous closed canopy and only few and small gaps (Guariguata & Ostertag, 2001). This is in contrast to the more conventional forest structure indicator of mean canopy height, where there was a clear progression from recently logged, regrowing, and old-growth forests (Figure 1 and Table 1).

Our results add another dimension to the often nonlinear recovery trajectories that characterize secondary succession in tropical forest (Norden et al., 2015) and the complex boom-and-bust dynamics after disturbance observed in plant communities more generally (Stott et al., 2010). They also suggest that regenerating logged forests may require a specific framework that explicitly addresses their distinct structure and functioning and contribution to the global carbon cycle (Bongers et al., 2015). In this respect, remote sensing technologies such as ALS, as well as satellite-based laser scanners like GEDI (Dubayah et al., 2020), have the potential to transform our understanding of how natural and human disturbances that shape the three-dimensional structure of forests in different ecosystems. But to realize the full potential of these remote sensing technologies, we may need to re-examine what structural attributes we derive from the data and what ecological inferences they allow us to make.

AUTHOR CONTRIBUTIONS

TJ conceived the idea for the study. DAC coordinated the airborne laser scanning survey of the SAFE project. BZ led the analysis and writing, with assistance from FJF and TJ. All authors contributed substantially to revisions.

ACKNOWLEDGMENTS

We thank NERC's Airborne Research Facility and Data Analysis Node for conducting the airborne laser scanning (ALS) survey and Tom Swinfield for help with processing the ALS point cloud data. We acknowledge the Sabah Biodiversity Centre, Sabah Biodiversity Council, Maliau Basin and Danum Valley Management Committees and the Economic Planning Unit for their support, access to field sites and for permission to carry out the ALS surveys in Sabah. We also wish to thank the South East Asia Rainforest Research Partnership (SEARRP), Sabah Foundation, Benta Wawasan, the State Forest, Sabah Chief Minister's Departments, and the Sabah Forestry Department. Finally, we are extremely grateful to Wan Shafrina Wan Mohd Jaafar for translating the abstract of the paper into Malay. The acquisition of the ALS data was funded through a grant awarded to the BALI consortium through NERC's Human Modified Tropical Forests Programme (grant code: NE/K016377/1). TJ was supported by a NERC Independent Research Fellowship (grant code: NE/S01537X/1) and through a Research Project Grant from the Leverhulme Trust which also funded FJF (grant code: RPG-2020-341). BZ was funded by a Scholarship from the China Scholarship Council (grant code: 202008320276).

CONFLICT OF INTEREST

The corresponding author confirms, on behalf of all authors, that there have been no involvements that might raise the question of bias in the work reported or in the conclusions, implications, or opinions stated.

DATA AVAILABILITY STATEMENT

All ALS-derived canopy height models used in this study are publicly archived on the SAFE project’s Zenodo data repository: https://zenodo.org/record/4020697#.YgFHmN_P2UK. Data and R code needed to replicate the analyses presented in this manuscript are archived on the following Zenodo data repository: https://doi.org/10.5281/zenodo.7339041.

ORCID

Tommaso Jucker https://orcid.org/0000-0002-0751-6312

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

---

**How to cite this article:** Zhang, B., Fischer, F. J., Coomes, D. A., & Jucker, T. (2022). Logging leaves a fingerprint on the number, size, spatial configuration and geometry of tropical forest canopy gaps. *Biotropica*, 00, 1–14. [https://doi.org/10.1111/btp.13190](https://doi.org/10.1111/btp.13190)