Preprint bg-2022-251

Response to Reviewer 1

The paper is an interesting and valuable contribution to the literature. Its unique value lies in connecting field-based measurements with remote sensing measurements, and then using this approach to link gap dynamics with biomass losses. As pointed out in previous reviews, the main underlying issue is one of sample size, but the authors have (mostly) addressed this now in their analysis and discussion. Similarly, earlier points on gap definitions precipitation have been addressed. I also highly value the authors' efforts in compiling and analysing the data as well as the detailed responses.

However, there are still two main issues that need to be addressed.

Interpretation/discussion of results: I found the interpretation of the results and the discussion sections difficult to follow and sometimes misleading. The clearest instance of a misleading claim was Discussion section 4.4, which states "Extreme rainfall-events control gap formation". This is a causal claim that is astonishing given both the study's setup (small spatial and temporal scales) and the study's results (at best weak correlations between gap patterns and precipitation). Similarly, section 4.3 states "Small-scale disturbances dominate canopy dynamics and associated biomass losses in Central Amazonia". This was also surprising, because the authors never correlate gap size and biomass losses (which they absolutely should, cf. below!). From the available materials, it seems to me rather that tree snapping, i.e., the mechanism responsible for the largest gaps, accounted for the largest biomass losses, so the opposite of what the paragraph claims. I may be mistaken, but then the authors need to show this. In contrast, sections 4.1 and 4.2, while not misleading, seem to summarize a wide variety of issues without a coherent storyline. Some of the points, such as detection sensitivity in the first paragraph need to be explained, other points should be summarized much more so that readers have clear takeaways. Part of my confusion may be due to the authors attributing properties of methods to the forest ecosystem itself. E.g., in stating "4.1 The mechanism of gap formation is related to the sensitivity of detection" they likely mean that the "classification of gap formation mechanisms" is related to detection sensitivity, not the actual mechanism. Another aspect may be that the authors tend to focus a lot on p-values/significance both in discussion/results, but not on effect sizes (e.g. how large are differences), and sometimes report with a precision that is not warranted for the small sample size (e.g. differences between 6, 7 or 8 gaps between mortality modes are not important, but seem more important when rendered as 2-digit precision percentages). I have made suggestions in the detailed comments to improve this.

The last version of our manuscript already included clear statements on the limitation of our results due to the relatively low number of gaps. However, we would like to stress that our monitoring extended for more than two years and the sampling size also reflects the tree mortality rates of the forest ecosystem we have worked in. Still, we have revised the entire text again to address all points raised by reviewer 1 in the Discussion section of our revised manuscript.

Biomass: I thank the authors for putting the overall biomass losses into the context of the carbon stocks at the site, but I still find that the analysis does not fulfill its potential here. I strongly recommend adding one or two scatterplots/correlation plots to Figure 6. This would be one plot c) that plots biomass losses against gap size, and one plot d) that plots biomass losses against gap perimeter or GSCI. In each case, the data points should be coloured by gap formation mechanism. That would quickly tell readers how easy (or difficult) it is to infer biomass losses from gap formation, and would provide an interesting result + directly relate to Discussion Section 4.3.

We thank the reviewer 1 for this suggestion and replaced Figure S3 by these scatter plots:

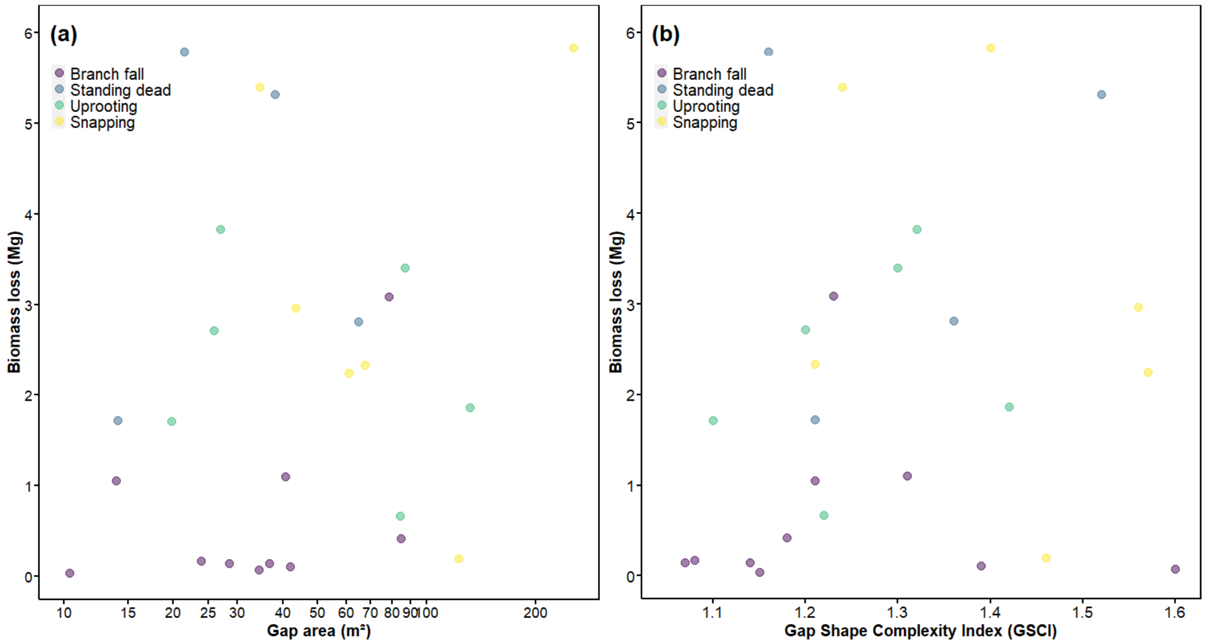

Figure S3: Relationship between biomass loss and gap area (a) and between biomass loss and Gap Shape Complexity Index (GSCI) (b) at the INVENTA plot, Central Amazon, Brazil. The data were collected from September 18th 2018 to January 19th 2021. Gap area was determined from UAV Imagery data. X-axis in panel *a* is log-scaled.

Finally, I appreciate that it is hard to track all changes in revisions, but there are a quite a few paragraphs where words or descriptions are missing or where the authors have only responded to the reviewers, but not included the respective text in the article (e.g., log-transformation of variables, when performed, needs to be mentioned in the methods).

Detailed comments:

Title: “in a Central Amazon forest”?

R: We corrected the word in the title.

l.14: “In addition to detecting” could be removed, i.e., just write: “We measured the size …”

R: The text was removed.

l.17: “corresponding” instead of “associated”?

R: We took the suggestion.

l. 17: “either … or” not “either … and”

R: The text was correct.

l. 22: “Regardless of”

R: The text was corrected.

l. 23: “lognormal” should probably be lower case, as it is not a name (also in the rest of the text)

R: We incorporated the suggestion throughout the text.

l. 22-26: not sure whether I would put the different fits of lognormal/Weibull into the abstract. As stated in earlier reviews, given the small sample size such fitting exercises should not be over-emphasized, especially when there is no clear theoretical justification why one should be better than the other.

R: We suggest that: "Regardless of the detection method, the size distribution was best described by a lognormal function for gaps starting from the smallest detected size (9 m2 and 10 m2 for field and imagery data, respectively), and the Weibull and Power functions for gaps larger than 25 m2. Properly assessing associated confidence intervals requires larger sample sizes. "

l. 26: I would reframe: “The main modes of tree mortality were not related to gap size, but to losses in biomass.” The problem is, however: what does this have to do with gap patterns? Can we detect differences in mortality/biomass losses from the gaps themselves (through size/perimeter)? This is one of the main issues referred to above.

R: We believe that it is possible to distinguish the mechanisms of gap formation by combining our UAV images and machine learning technics. However, this would require complementary (and extensive) fieldwork, and it was not within the scope of this study. We keep monitoring the plot and are currently working on these aspects, which will be addressed in an upcoming article.

l. 27: Again, not sure whether this is such a strong result.

R: We made clear that our results shall be carefully interpreted due to the relatively low number of gaps. Extrapolations of observed patterns to other Amazon regions also require validation of our method and further data sampling. However, as mentioned in the previous round, the significant association with extreme events aligns with what has been previously reported for the same region (Negrón-Juárez et al., 2017; Fontes et al., 2018; Aleixo et al., 2019). *We have improved the text to make limitations clear and while suggesting issues that shall be addressed in further studies.*

l. 28: Does this not contradict l.26?

R: Our goal here was to highlight the overall contribution of gaps. We worked again in the text to clarify that gaps ≥ 1 ha are not uncommon and as previously shown, can be traced using larger scale optical satellites (e.g., Sentinel-2, Landsat) (Marra et al. 2014; Negron-Juarez et al. 2011; Emmert et al. in review (doi: 10.20944/preprints202305.1631.v1).

l. 32: This sentence is too harsh in my opinion. It’s not like these results don’t tell us anything about biomass/gap dynamics, it’s just that we have to be careful about extrapolating. I would drop the sentence about extrapolation, or merge it with the following: “Future investigations combining remote sensing with field data are needed to confirm these relationships at landscape scale.”

R: The sentence was changed as following: "While combining remote sensing with field data has proven to be an accurate and precise method for mapping gaps compared to other existing approaches, it is important to note that our sample size is still relatively small. Therefore, the extrapolation of the results beyond our study region and landscape shall be made cautiously."

l. 53: Maybe combine the sentences: "The size of gaps can vary … and defines the amount of light …"

R: We combined the two sentences.

l. 54: "Apart from related" – there is something missing

R: The text was changed to: "In addition to being influenced by mechanisms of formation, the size and shape of gaps can also be influenced by local climate, extreme weather events, topography, soil, forest structure and species composition (Denslow, 1987; Marra et al. 2014; Araujo et al., 2021; Cushman et al., 2022)."

l.58: Remove "Although with some surviving trees"

R: The text was removed.

l.57-70: I would merge these lines into one paragraph on wind and rain, and link them to the previous paragraph, e.g. by starting after "and related functions (Jucker, 2022)" with "In the Amazon, one of the key disturbances are extreme wind and rainfall events…." The following sentences need a bit of reworking/reordering.

R: The paragraph was modified as following:

"Extreme wind and rain are major mechanisms of tree damage and mortality in Amazon forests (Nelson et al., 1994; Chambers et al., 2013; Magnabosco Marra et al., 2018; Urquiza Muñoz et al., 2021; Negrón-Juárez et al., 2023). Gaps opened by extreme wind and rain can have areas greater than 3,000 hectares (Nelson et al., 1994; Espírito-Santo et al., 2014; Negrón-Juárez et al., 2010, 2018, 2023). Recent studies have identified windthrow hotspots (Negrón-Juárez et al., 2023), with their frequency influenced by climate-sensitive atmospheric phenomena, such as convective potential energy (CAPE) (Feng et al., 2023). Previous studies reported higher frequency of gaps during wet months, even in drought years, which suggests a positive correlation between precipitation and tree mortality (Fontes et al., 2018; Aleixo et al., 2019). Additionally, satellite data support that large-scale windthrows (Magnabosco Marra et al., 2014b) visible on Landsat imagery (Negrón-Juárez et al., 2011) occurred more frequently between September and February, which are months marked by extreme rainfall (eg., > 30 mm hour-1; Negrón-Juárez et al., 2010, 2017)."

l. 74-75: maybe make it a bit simpler "from small number of plots and infrequent surveys is a challenging task"

R: The sentence was rewritten.

l.76-77: I would focus only on optical methods here or rewrite this and the following with a clear separation of lidar/optical or satellite/airborne, otherwise it gets confusing for the reader. I.e., first you mention lidar, e.g. as in Greg Asner's studies, but then you move on to Landsat, which operates in a very different realm, both in terms of resolution and data type. Also, as an aside, while I agree that airborne lidar generally provides high resolution and accuracy for e.g., canopy height, I would be more careful about this statement in the context of gap dynamics. Gap patterns depend a lot on CHM generation algorithms and the resulting surface roughness, so accuracy is probably not well-defined.

R: We removed the sentences. The revised text starts with "In the Amazon, studies using intermediate spatial-resolution..."

l.88: Again, it makes much more sense to only focus on optical data in the previous paragraph if the focus is on optical data here. Otherwise, a comparison with lidar would be needed. E.g., optical UAV are probably much cheaper to acquire and easier to handle, so practical (but they probably also come with a few disadvantages, such as not penetrating to the ground, being more sensitive to low clouds/mist).

R: We changed the previous paragraph as suggested.

l.114: This is not a crucial comment and I am not familiar with the INVENTA plot, but out of interest: this seems an extremely narrow height range for canopy trees (only ~50cm standard deviation in height?). From most height allometries at plot level that I know, tree height variation for trees > 50cm DBH is usually massive (several meters). Are canopy trees very narrowly defined here?

R: This analysis was conducted with 267 canopy trees with DBH > 25 cm along a 1 km transect encompassing plateau, slope, and lowland areas. The analyses of tree proportion, diameter distribution, and growth rates between canopy and understory, as reported in Araujo, 2019 and Araujo et al., 2020, *support the finding that the 267 trees analysed have reached the equilibrium phase of the successional stage and therefore have similar heights. For instance: (i) growth is not related to diameter for canopy trees and positively related for understory trees, consistent with the idea that light availability increases with diameter for understory trees but not for canopy trees; (ii) for canopy trees, the exponential distribution, typical of forests in equilibrium condition, is only fitted to trees with DBH > 25 cm; (iii) trees with DBH > 25 cm have more than a 50% probability of being in the canopy.*

l.198: could not find Magnabosco Marra et al. 2016 in references (was it 2014?)

R: This reference was indeed missing in the previous version and is now included.

l.230: "three" should be "four" now

R: The suggestion was incorporated.

l.245: "two-tailed" what? In your responses you mention the log-transformations, but this needs to be also part of the methods!

R: We thank reviewer 1 for pointing this issue. We included the logarithmic transformations as part of the Methods section.

l.268-289: This goes back to my questions about your gap definition in the previous round. It is important to differentiate here between what one of the methods registers (UAV) and what has happened in the actual forest (as seen, e.g., from all combined methods). Since the field measurements found three gaps that fulfill Brokaw's definition, it's not correct to say that "there was no traceable change in the upper canopy of the forest". There clearly was a traceable change, it was just not traceable by UAV. This should be carefully phrased throughout the article.

R: We changed to: "These gaps, which were not visible on either the difference images or the orthomosaics, indicate an absence of detectable change in the upper canopy."

l. 274-277: I would rephrase this (cf. also my comments from last time). A non-significant p-value is not the same as a small effect size (e.g., “no evidence for strong differences”). I would suggest: “UAV-inferred gap size exceeded the field-inferred gap size substantially for the largest gap (254 m2), and on average, the difference was 11.5 m2 (17% of mean field gap size), but non-significant (p = 0.85).”

R: We changed the text as suggested.

l.277-278: Make sure to check whether normal distributions actually are appropriate for those metrics as well, or if data were log-transformed or otherwise transformed, this needs to be part of the methods/tests descriptions. If possible, focus on effect sizes.

R: The perimeter and shape complexity index (GSCI) data are log-transformed. We included this information in the descriptions of the Methods section.

l.289: I do not understand what “The two most discrepant” refers to here

R: These are the two lowest values of mean canopy height loss (Fig. 4). We changed the sentence to: "The two gaps with the lowest values of height loss (i.e., 1.13 m and 2.13 m) were only detected in the field (Fig. 4).”

l.303-305: Please provide the number of gaps used for each analysis, and maybe use “best described” instead of “better described”. As per the more general comments, it’s fine to keep these results, but I would not put too much emphasis on them.

R: We included the number of gaps used for each analysis and modified to the text as "best described".

l.307-313: I really like this analysis, but I would not overemphasize the differences gap numbers. Differences between 8,7,6 (and even 11) gaps are hardly important. Also, I would not report percentages in Table 2, or at least not with 2-digit precision! With 32 gaps, a single gap results in a change of ca. 3%, so a precision < 1% is meaningless. You can always summarize as 34%, 25%, 22% and 19%, but probably 35%, 25%, 20% and 20% would be the most appropriate precision. I would write: “All mechanisms of gap formation accounted for a substantial share of the gaps created (from ~20% for standing dead trees to ~35% for branch falls). In contrast, contribution to total gap area was highly asymmetric, with ~60% accounted for by tree snapping, and <10% by standing dead trees.

R: We rewrote the paragraph as suggested.

l.317-321: As before, I would focus much less on the p-values. P-values usually tell us very little (for large sample sizes they are always < 0.05, for small samples, very rarely). How about effect sizes? Why not simply describe what you see in Figure 6a and 6b? From the graph it is clear that most gaps are very similar in size, irrespective of gap formation mode, but that snapping in this case has much more variance. So I would describe that. And from 6b, we see that all three tree fall modes have higher biomass loss than branch-fall, which also makes a lot of sense, so I would describe this! It would be important to add one or two panels showing the correlation between gap area (or gap perimeter) and biomass loss as a scatter plot. I.e., x axis is gap area/gap perimeter, y axis is biomass loss, and dots are coloured by mortality mode. I think what most people would be interested in is: if I measure the gap area and gap perimeter, how well can I predict the biomass loss?

R: We reduced the emphasis on p-values and instead focused on the effects’ size. We have also included the requested figure (please, see response to the overall comment).

348: The header for this whole paragraph is a bit odd, because this point (sensitivity of detection, mechanism of gap formation) is barely explained and probably factually wrong. The mechanism of gap formation itself is likely not dependent on detection, because it is a property of the forest ecosystem itself. I assume that the authors mean "The classification of gap formation methods is related to detection sensitivity". But again, this is also not explained, and the section as a whole lacks in coherence.

R: The text was corrected as: "Detection is influenced by modes of tree mortality, branch fall and consequent gap features "

348-350: By definition, the approach does not only underestimate the frequency smaller than the size threshold, it does not quantify it at all, so I would remove this.

R: Removed.

351=354: The caveat about needing more studies across different landscapes might make more sense later in the discussion, but not an urgent alteration.

R: We sincerely thank you for the review. It is of utmost importance to emphasize the need for conducting studies across various landscapes. This aspect has been addressed throughout the manuscript to reinforce the accuracy of our data, avoiding any misleading interpretations.

357-358: This needs to be explained more.

R: We added to the text: "It is important to note that the sensitivity for detecting gaps was influenced by modes of tree mortality, which also lead to specific effects on species composition in regenerating patches of forest (Putz et al., 1983; Chao et al., 2009)."

385: Again, the header is difficult to understand and does not summarize all the points that are discussed in the paragraph (GSCI, power law vs. lognormal). I would shorten these paragraphs a lot and clearly say what is essential. E.g., lines 425-430 seem to describe how the lognormal function is fitting the data, but it is very hard to follow. Also, as stated before, I would not focus too much on these results. I would just state the simple things: Lognormal was a bit better. This may reflect underlying processes, but should not be overinterpreted due to methods and small sample size. I don't think the results allow for much more inference.

R: We changed the text to "Gap geometry and size structure: differences between imagery and field data". We also reduced the emphasis on model fitting and removed lines 425-430.

431: This header also does not make much sense – is this really what your data show? A scatterplot of biomass loss vs. gap size would go a long way in explaining this, so I strongly recommend to make this clearer both in Results and Discussion! If I look at Figure 6b, some of the strongest biomass losses are recorded for tree snapping, whereas branch losses predictably have a lot of very small biomass losses.

R: We changed the header to: "Ecosystem importance of mechanisms of gap formation and associated losses of tree biomass"

431-448: The whole paragraph seems to summarize more the literature than relate to the results.

R: We modified the text as suggested and included the following paragraphs for discussing our results:

"Repeated field measurements allowed us to quantify the relative importance of modes of tree mortality (i.e., standing dead, snapping and uprooting) and branch fall. Our results show that biomass losses did not differ among mechanisms of gap formation, but was relatively smaller in more frequent gaps formed by branch fall. Nonetheless, we could not distinguish different of gap formation based on the gap area and associated losses of biomass. These unclear differences indicate that quantifying the importance of varying mechanisms of gap formation (referring to the vectors that cause damage or death to trees and, consequently, gap formation) and modes of tree mortality (referring to mortality forms such as standing

dead, snapped, and uprooted trees) is a challenging task that require larger sample sizes and imagery data with higher spatial and spectral resolution.

Our findings corroborate those from a previous study conducted in Santarém (also in the Brazilian Amazon) using repeated high-density LiDAR data (Leitold et al., 2018). This study revealed that biomass losses due to single and multiple branch fall events accounted for only 20 % of the estimated biomass loss from canopy and understory trees. Similarly, in Panama, branch fall was associated with 43.5 % of gaps formed over a five-year period, but only for 23 % of the total disturbed area (Araujo et al., 2021b). The consistency of these findings across different tropical forests highlights the importance of tree mortality and canopy structure for regulating biomass stocks and balance.

To our knowledge, this is the first study combining remote gap-detection with direct measures of biomass (i.e., branches) and estimates based on a locally adjusted allometry. Almost half of the aboveground biomass of tropical forests (42 %, range of 12 % – 76 % across forests) is lost due to damage to live trees (Zuleta et al., 2023). If climate change results in a higher frequency of storms and extreme winds (Feng et al., 2023; Negrón-Juárez et al., 2023), branch fall and tree mortality rates can also be expected to increase. This may affect carbon stocks and balance, as well as the functional composition of these forests at the landscape level (Magnabosco Marra et al., 2018; Urquiza Muñoz et al., 2021)."

449: The final header seems misleading. First, as outlined before, the sample size of the data makes this analysis problematic, but more importantly, if I understand the results correctly, there were no strong correlations in the data, so how do we come to the conclusion that "Extreme rainfall-events control gap formation"?

R: We changed to: "Extreme wind and rainfall as potential mechanism of gap formation".

474-475: As stated in your abstract and in line with previous comments on sample size, the study cannot "reliably assess landscape patterns". It can only assess local patterns and indicate what might hold elsewhere.

R: Although at the local level, our study address landscape variations of topography and soil, including areas of plateau, slope and valleys. Since the scale of the study is clearly stated in the Abstract and the Material and Method section, we edited the text as following: "By combining high temporal and spatial resolution UAV imagery with detailed field data integrating landscape variations of topography and soil, our study provides novel and fundamental knowledge for understanding how tree mortality processes affect the structure and dynamics of Amazon forests."

475: "Mechanisms of gap formation could only be distinguished in the field." To my knowledge this has not been shown in the study. If this is an important result, then it needs to be shown, e.g. by plotting not only gap area against gap formation mechanism, but all gap properties.

R: This was supported by our data and as clearly stated in the text, refers exclusively to our study site and applied dataset. Nonetheless, we addressed the importance of this aspect in the Discussion section and pointed out that UAV images can be further explored to this goal. Other arguments complementary to this aspect are also provided in our answer to the comment in line #26.

Response to Reviewer 2

This updated version of the manuscript is in much better shape and so I recommend the manuscript to be accepted as is. After reading the author's responses to my comments and reading the revised version, I got a better understanding of the number of gaps being mapped by either field and UAV.

The beginning of the Discussion highlights the excellent results for the match between field x UAV, then authors added a sentence acknowledging that more studies are required to further confirm these findings, because in the end we are taking an n=18 gaps with field + UAV to reach Conclusions. While the calculated F1-Score is excellent and trully above any other method that exists to map gaps, sample size is still very small to be sure this is a robust method. If I may still suggest something to be added, up to the authors, I think the manuscript would benefit (and I would be very much interested into reading about) the authors' perspective on a few sentences about potentials and limitations on scaling these maps of gaps to larger scales, that is, how to approach this in terms of data (sentinel-2, planetscope) and measuring gaps in the field for cal/val (recommendations for how to measure them in the field "quickly"), etc. Nevertheless, good job on the study and on undertaking the challenge to match field with remote sensing. I look forward to reading more manuscripts from this great team. Att, Ricardo Dalagnol.

R: We appreciate the constructive and positive comments provided by reviewer 2. Following his suggestions, we included a detailed discussion on the limitations and opportunities of scaling up UAV maps to larger scales, such as satellites, and field measurements for calibration and validation of methods of gap detection:

"Scaling down is the first step to scale up processes and mechanisms regulating forest dynamics and functioning. This requires robust and validated remote sensing tools and method integrating regional variability of forest and environmental attributes. However, optical sensors with wide spatial coverage, relatively short revisiting time, and long data series are still limited in the Amazon. For example, in Landsat images, mortality events involving clusters with fewer than 6-8 trees cannot be identified (Negrón-Juárez et al., 2011; Chambers et al., 2013). Furthermore, the smallest gap size found in our study was 10 m² using a 1 m² elevation model and 2 cm resolution orthomosaics. The spatial resolution of Sentinel-2 is 10 m, while that of Planet is 3 m to 5 m. Therefore, the smallest gap size would correspond to just 1 pixel in Sentinel-2, potentially resulting in possible underestimation of small gaps. Extensive field inventory provides even more valuable information for scaling down (Fontes et al., 2018). However, it is necessary to develop pre-established field protocols that are reproducible and functional. These limitations can be mitigated by expanding drone-imaging coverage areas and increasing spectral information on targets, which should be the focus of future work. In conjunction with geometric patterns of gaps, as described in our study, higher spectral resolution can contribute not only to accurately distinguishing tree mortality modes and branch fall, but also to improving landscape estimates of biomass loss and recovery. Larger drones with the capacity to carry sensors designed to collect data across a broader spectrum range are fundamental for enhancing existing methods and establishing routines that enable detailed assessments of canopy dynamics and associated processes in dense and diverse tropical forests."