Response referee #2

This manuscript investigates evapotranspiration (ET) and evaporative fraction (EF) across 52 experimental tree islands in an oil-palm landscape using UAV thermography, the DATTUTDUT energy-balance model, random-partitions analysis, and piecewise structural equation modelling. The study addresses a relevant gap by quantifying ET and EF across the full tree-island area and planted-diversity gradients of the EFForTS-BEE experiment.

The manuscript is generally well structured and presents a high-quality study.

I identified some aspects for which additional clarification or sensitivity analyses would strengthen the interpretation. These principally concern the selection of one observation per island, the partial confounding of island area with palm removal, the proposed edge-based explanation, and the reporting of direct and indirect SEM effects. Overall, the positive relationship between island area and EF appears to be an important contribution. Clarifying the robustness of this relationship, its separation from the establishment treatment, and the evidence for the proposed underlying mechanisms would further improve the manuscript.

Major Comments

1. Selection of one observation per island

Selecting the observation with the highest incoming radiation is an understandable approach to standardize measurement conditions, particularly because 98 % of the selected observations fall between 850 and 1000 W m-2. However, comparable radiation does not necessarily imply comparable soil-water availability, vapour-pressure deficit, wind conditions, or antecedent rainfall.

> We agree that comparable incoming radiation does not necessarily imply comparable environmental conditions. To assess this, we compared the six high-radiation flight days selected for modelling with the other eight usable flight days during the campaign period using data from a climate station located next to the experiment. Soil moisture was very similar between the two groups (32.88% [IQR 32.64–33.00] vs. 33.19% [32.71–33.68]), as were antecedent precipitation over the preceding 1 day (0.05 vs. 0.25 mm) and 3 days (3.35 vs. 3.45 mm). The selected days had lower 7-day antecedent precipitation (8.20 vs. 57.95 mm), while daytime air temperature was somewhat higher (32.54 vs. 31.28 °C) and wind speed slightly lower (0.60 vs. 0.69 m s^{-1}). Relative humidity and VPD could not be assessed reliably because the humidity sensor was unavailable during part of the campaign. Overall, the selected flight days experienced broadly comparable soil-moisture, precipitation, and wind conditions to the other usable flight days.
>
> We re-structured section 2.4 (Methods) accordingly:
>
> 98% of these selected observations (n = 52) were obtained under high incoming shortwave radiation (850 - 1000 W m^{-2}). The selected flight days showed broadly comparable soil-moisture, antecedent-precipitation, and wind conditions to the other usable flight days, as measured by a climate station in the oil palm plantation (Supplementary materials, table S1). In contrast to ET, EF is only marginally affected by varying levels of radiation (Fig. A1). [240-244]

Table S1: Environmental conditions on selected flight days (high incoming radiation) vs. other days with available data. Median [IQR] for soil moisture on flight days, aggregated precipitation for 1, 3 and 7 days before selected flight day, air temperature and wind speed on flight days.

Variable	Selected high-radiation flights (n = 6)	Other usable flight days (n = 8)
Soil moisture (%)	32.88 [32.64–33.00]	33.19 [32.71–33.68]
Precipitation preceding 1 d (mm)	0.05 [0.00–1.75]	0.25 [0.00–2.20]
Precipitation preceding 3 d (mm)	3.35 [2.52–4.62]	3.45 [1.42–5.72]
Precipitation preceding 7 d (mm)	8.20 [8.20–43.90]	57.95 [44.30–72.10]
Air temperature, 10–14 UTC+7 (°C)	32.54 [31.70–32.79]	31.28 [30.98–31.96]
Wind speed, 10–14 UTC+7 ($m\ s^{-1}$)	0.60 [0.58–0.61]	0.69 [0.58–0.72]

A clearer explanation of the selection rule would help readers understand this analytical choice. This could include reporting within-island EF variability and comparing the reported area effect with an analysis based on the repeated observations or another suitable island-level summary. If the magnitude of the relationship is sensitive to the selection method, the reported 17% increase could be described more specifically as applying to the selected high-radiation observations.

We thank the reviewer for this suggestion. To assess whether the observed island-area effect was sensitive to our selection of the observation with the highest incoming shortwave radiation, we conducted a sensitivity analysis using all repeated EF observations available for each island. We retained the same 48 islands and the same random-partition modelling framework as in the original analysis, but, in each of 5,000 iterations, randomly selected one observation from each island. Thus, removing the dependence on the maximum-radiation selection rule.

The resulting island-area effect was highly robust to this alternative selection procedure. Island area remained statistically significant ($p < 0.05$) in 99.4% of the 5,000 iterations; the median p-value for island area was 0.00015. Furthermore, the raw EF increase between the smallest (25 m^2) and largest (1600 m^2) island areas was 18.5% on average across the random selections (95% range: 11.22–26.41%), compared with the 17% increase obtained from the original high-radiation observations. Thus, the magnitude and direction of the area effect were very similar when observations were selected independently of incoming radiation.

We therefore conclude that the reported relationship between island area and EF is not driven by the selection of the highest-radiation observation. We have added this sensitivity analysis to the revised manuscript and clarified the observation-selection procedure accordingly.

We re – structured section 2.4 (Methods):

The number of observations per tree island ranged from 5 to 10. To standardize measurement conditions, we selected, for each tree island, the observation recorded under the highest incoming shortwave radiation. This selection was made before analysing the factors controlling EF and was intended to minimize variation in radiation among observations. The selected observation was therefore not necessarily representative of the mean EF of each tree island during the measurement campaign. 98% of these selected observations (n = 52) were obtained under high incoming shortwave radiation (850 - 1000 W m^{-2}). [239-244]

And section 3.1 (Results):

From the smallest island area (25 m^2) to the largest (1600 m^2), EF increased by 17%, from an average of 0.65 to 0.76. To test whether the observed island-area effect depended on selecting the observation with the highest incoming shortwave radiation, we conducted a sensitivity analysis using all EF observations available for the 48 tree islands. In each of 5,000 iterations, one observation was randomly selected from each island, and the original random-partition model was refitted. Island area remained significant ($p < 0.05$) in 99.4% of the iterations, with a median p-value of 0.00015. The mean raw EF increase between the 25 and 1600 m^2 islands was 18.5% across the random selections (95% range: 11.22–26.41%), compared with 17% based on the selected high-radiation observations. Thus, both the direction and magnitude of the area relationship were similar when observations were selected independently of incoming radiation. [332-339]

2. Island area is partially confounded with palm removal

The 25 m^2 islands were established without removing oil palms, whereas 40 % of the palms were removed from the 100, 400, and 1600 m^2 islands. The smallest area class therefore differs from the larger classes in both island area and establishment treatment. These differences could include residual palm density, light availability, competitive conditions, and potentially soil disturbance and water availability.

Because palm removal is tied to area class, its contribution cannot be fully separated simply by adding it as another predictor. I suggest acknowledging this feature of the experimental design more explicitly. If supported by the available sample size, a useful sensitivity analysis would be to assess the area relationship using only the 100, 400, and 1600 m^2 classes, which received the same nominal palm-removal treatment. If the positive relationship remains evident among these three classes, this would strengthen its interpretation as an area effect. If it depends mainly on the contrast with the 25 m^2 class, the interpretation should acknowledge that island area and establishment treatment cannot be clearly separated for this particular comparison.

We thank the reviewer for highlighting this important feature of the experimental design. We agree that palm removal is intrinsically linked to area class and therefore cannot be fully separated from the area effect in comparisons involving the 25 m^2 islands.

This design reflects a deliberate compromise during the setup of the EFForTS-BEE experiment. The smallest tree island class was established between oil palm rows as requested by local practitioners (no cutting of oil palms).

To assess how much the positive area–EF relationship depends on the contrast with the uniquely treated 25 m² class, we conducted a sensitivity analysis restricted to the 36 islands in the 100, 400, and 1600 m² classes, all of which received the same nominal 40 % palm-removal treatment. We randomly selected one of the repeated EF observations for each island and repeated the original random-partition model in 5,000 iterations. The positive relationship between island area and EF remained evident after excluding the 25 m² class. Across the random selections, mean EF increased monotonically from 0.68 in the 100 m² islands to 0.71 in the 400 m² islands and 0.74 in the 1600 m² islands. This ordering of the raw EF means (100 < 400 < 1600 m²) occurred in 92.1 % of the random-selection iterations. The corresponding ordering of the adjusted island-area coefficients in the random-partition model occurred in 91.2 % of iterations. The mean increase in raw EF between 100 and 1600 m² was 8.8 % (95 % sensitivity range: 3.0–14.7 %).

Statistical support for the overall island-area effect was reduced compared with the full analysis, with island area remaining significant ($p < 0.05$) in 57.2 % of the random-selection iterations (median $p = 0.0347$). We interpret this difference as indicating that, although the statistical strength of the area effect is reduced when the 25 m² class is excluded, the direction of the relationship remains highly consistent, at a somewhat lower effect size, across the three area classes receiving the same palm-removal treatment.

Together, these results indicate that the positive association between island area and EF is not solely driven by the contrast involving the uniquely treated 25 m² islands. At the same time, the stronger 17 % increase observed between 25 and 1600 m² cannot be interpreted as a pure area effect independently of establishment treatment. We have therefore revised the manuscript to acknowledge explicitly that area and palm-removal treatment are confounded for the 25 m² versus larger-island comparison, while noting that the positive area relationship remains evident among the three area classes that shared the same palm-removal treatment.

We re – vised section 2.1 (Methods):

Before tree planting 40% of oil palms were removed inside of the 100 m^2, 400 m^2 and 1600 m^2 plots to increase light availability (Teuscher et al., 2016). In contrast, the 25 m^2 plots were established between oil – palm rows, and no oil palms were removed. This design reflects a compromise between increasing light availability for the planted trees and retaining the prevailing oil – palm management system. Consequently, the 25 m^2 class differs from the larger area classes in both island area and establishment treatment. In addition to tree planting, the experiment also allowed for natural plant recruitment. [180-184]

Section 3.1 (Results):

Because palm removal differed between the 25 m² islands and the three larger area classes, we conducted a sensitivity analysis using only the 36 islands in the 100, 400, and 1600 m² classes, which all received the same nominal 40% palm-removal treatment. In each of 5,000 iterations, one repeated EF observation was randomly selected for each island, and the original random-partition model was refitted. Mean EF increased monotonically from 0.68 in the 100 m² islands to 0.71 in the 400 m² islands and 0.74 in the 1600 m² islands. The ordering of the raw means (100 < 400 < 1600 m²) occurred in 92.1% of iterations. The mean increase in raw EF between the 100 and 1600 m² islands was 8.8% (95% sensitivity range: 3.0–14.7%). Statistical support for the overall island-area effect was weaker than in the full analysis: island area was significant ($p<0.05$) in 57.2% of iterations, with a median p-value of 0.0347. Therefore, although excluding the 25 m² class reduced statistical support for the overall area effect, the direction of the relationship the size of the effect remained consistent among the three area classes that shared the same palm-removal treatment. [356-365]

Section 4.2 (Discussion):

Interpretation of this area effect requires consideration of the experimental design. The 25 m² islands were established without removing oil palms, whereas 40% of the palms were removed from the 100, 400, and 1600 m² islands. Thus, the 17% increase between the 25 and 1600 m² islands cannot be attributed to island area independently of the establishment treatment. However, a sensitivity analysis restricted to the three larger area classes, which received the same nominal palm-removal treatment, showed a consistent and monotonical increase in EF with island area: mean EF increased by 8.8% between the 100 and 1600 m² islands. Although statistical support was weaker than in the full analysis, the direction of the relationship was highly consistent, indicating that the positive area relationship was not solely driven by the contrast with the uniquely treated 25 m² islands. [415-422]

3. Edge explanation

The manuscript reports higher EF near the edges of the 400 and 1600 m² islands. Under comparable island geometry and edge-effect width, however, smaller islands would be expected to contain a larger proportion of edge-proximate area. A higher EF near edges might therefore raise plot-mean EF more strongly in smaller islands, whereas the observed relationship shows increasing mean EF with island area.

It would be helpful if the authors could explain quantitatively how the reported distance-to-edge gradient relates to the observed increase in plot-mean EF with island area. This could include a clearer definition of the edge zone, its estimated spatial extent, and the proportion of each island represented by different distance-to-edge classes. The principal methods and results of the analyses referenced in Supplementary Tables S2 and S3 should also be summarized sufficiently in the main manuscript to make this interpretation traceable. A concise analysis or conceptual explanation connecting the within-island edge gradient to the among-island means would considerably strengthen the proposed mechanism. Until this relationship is demonstrated more clearly, I suggest moderating the statement in the Abstract that the area effect “can be explained” by the edge gradient. A formulation such as “may be related to the observed edge gradient” would better reflect the current evidence.

We thank the reviewer for pointing out this important geometric consideration. We agree that, because smaller square islands contain a greater proportion of edge-proximate area, a higher EF near island edges would not, by itself, be expected to explain the observed increase in mean EF with island area. Our spatial analysis showed an edge-associated decline in EF with increasing distance from the edge, particularly in the 20 × 20 and 40 × 40 m islands, with the transition occurring at approximately 30% of the island side length. However, the magnitude and spatial pattern of this gradient do not provide a sufficient explanation for the among-island increase in mean EF. Our SEM further identified significant indirect pathways from island area to EF through structural complexity (SSCI), both directly and sequentially via observed woody plant diversity and SSCI, indicating that vegetation structure provides an additional pathway linking island area to EF. A pixel-level analysis is ongoing, but not yet complete. Therefore, we have moderated the manuscript text and now describe the edge gradient as a potential contributor to, but not a sufficient explanation for, the island-area effect.

Re - vised Abstract:

The decrease in EF with increasing distance from the edge within larger tree islands suggests that edge-related spatial variation may be related to, but does not fully explain, the positive area effect. [54-56]

Re – vised section 4.2 (Discussion):

The direct area effect may partly reflect differences in EF between the edges of the tree islands and the surrounding oil-palm plantation. Our spatial analysis detected significant edge-associated declines in EF with increasing distance from the edge in the 400 and 1600 m^2 islands (Supplementary Table S3), with the transition occurring at approximately 30% of the island side length. However, because smaller square islands contain a greater proportion of edge-proximate area, this within-island gradient alone cannot explain the observed increase in mean EF with island area. We therefore interpret the edge gradient as evidence of spatial heterogeneity within the islands rather than as a sufficient explanation for the among-island area effect. [424-433]

We conclude that edges can induce changes in ET or EF, although the direction and magnitude of these effects vary among studies. This variability may reflect the origin of the forest patch, such as whether it is a remnant of old-growth forest with biomass loss at the edge (Ordway and Asner, 2020) or a recovering secondary forest, as well as the adjacent vegetation and energy advection, for example by wind. In our experiment, the observed edge-associated gradient indicates spatial variation in EF within the tree islands. However, because smaller islands contain a greater proportion of edge-proximate area, this gradient alone does not explain the increase in island-level mean EF with island area. The positive relationship between island area and EF may instead reflect a combination of edge-related and structural effects. [438-444]

4. Direct and indirect effects

The path coefficients reported in the Results differ from those shown in Figure 3. It would be helpful to clarify whether the text and figure present different effect metrics or standardization procedures and to ensure numerical consistency across the manuscript.

Figure 4 already presents bootstrapped direct, indirect, and total effects. Reporting the corresponding numerical estimates and confidence intervals in a table would make the relative contributions of the pathways easier to evaluate. In particular, it would be useful to distinguish:

- the positive serial pathway through observed woody plant diversity and structural complexity,
- the pathway from island area through structural complexity to EF, which opposes the positive serial pathway,
- the combined indirect effect, and
- the direct and total effects.

The conclusion states that island area enhances evaporative cooling "primarily" by increasing woody diversity and structural complexity. This wording does not appear fully aligned with the stronger direct area-to-EF relationship shown in the results. I therefore suggest distinguishing more clearly between the consistently detected overall relationship between island area and EF and the more model-dependent interpretation of the indirect pathways.

We thank the reviewer for highlighting the apparent discrepancy between the path coefficients reported in the Results and those shown in Figure 3. We re-examined the SEM and confirmed that the coefficients shown in Figure 3 are the standardized path coefficients (Std.Estimate) for the individual structural relationships in the SEM. Figure 4 and the corresponding Results text, in contrast, report bootstrapped direct, indirect, and total effects, which summarize the contribution of a predictor to EF across the network of direct and mediated pathways. Thus, the values in Figure 3 refer to individual arrows in the SEM, whereas the values in Figure 4 and the Results text refer to effects obtained by combining the relevant paths and quantifying their uncertainty through 1,000 bootstrap replicates. The apparent differences therefore reflect different effect metrics rather than inconsistencies in the SEM.

We agree that reporting the numerical estimates and confidence intervals for the individual indirect pathways makes the relative contributions of the pathways clearer. We therefore added Table S3, which reports the six individual indirect pathways from island area to EF, the combined indirect effect, and the direct and total effects, all with 95% BCa confidence intervals.

For island area, the direct effect on EF was 0.439 (95% BCa CI: 0.281 – 0.639), while the combined indirect effect was −0.044 (95% BCa CI: −0.213 – 0.103), resulting in a total effect of 0.395 (95% BCa CI: 0.314 – 0.532). Among the individual indirect pathways, the pathway from island area through observed woody plant diversity and structural complexity to EF was positive (0.045, 95% BCa CI: 0.010 – 0.130), whereas the pathway through structural complexity alone was negative (−0.066, 95% BCa CI: −0.185 – −0.014). The other indirect pathways were comparatively small and their confidence intervals included zero. Thus, although the positive serial pathway through observed woody plant diversity and structural complexity was supported, the opposing pathway through structural complexity resulted in a small and statistically unsupported combined indirect effect.

We have therefore revised the interpretation to distinguish more clearly between the consistently detected overall relationship between island area and EF and the more pathway-specific and model-dependent indirect mechanisms. In particular, we no longer state that island area enhances evaporative cooling “primarily” through increases in woody diversity and structural complexity. Instead, we describe the strong positive area–EF relationship as being predominantly represented by the direct area–EF pathway in this SEM, while noting that the positive serial pathway through observed woody plant diversity and structural complexity provides evidence for one potential mechanism contributing to this relationship.

*Table S3: Estimates are standardized effects from the SEM. Confidence intervals are 95% bias-corrected and accelerated (BCa) intervals based on nonparametric bootstrapping (*1000 iterations)*. Individual indirect effects are pathway-specific products of standardized path coefficients; the combined indirect effect is the sum of all indirect pathways.*

Pathway / effect	Estimate	95% BCa CI
Overall effects		
Combined indirect effect	-0.044	-0.213 to 0.103
Direct effect: island area → EF	0.439	0.281 to 0.639
Total effect of island area on EF	0.395	0.314 to 0.532
Individual indirect pathways		
Island area → structural complexity (SSCI) → EF	-0.066	-0.185 to -0.014
Island area → observed woody plant diversity → EF	-0.051	-0.229 to 0.087
Island area → GNDVI → EF	0.002	-0.019 to 0.053
Island area → height variability → EF	0.009	-0.010 to 0.059
Island area → observed woody plant diversity → structural complexity (SSCI) → EF	0.045	0.010 to 0.130
Island area → observed woody plant diversity → height variability → EF	0.016	-0.015 to 0.077

We re – vised section 2.9 (Methods):

Non-parametric bootstrapping with 1000 randomizations was used to calculate direct, indirect, and total effects, including confidence intervals for each predictor, using the “semEff” package version 0.6.1 (Murphy, 2022). Standardized path coefficients describe the individual structural relationships represented by the arrows in the SEM (Fig. 3), whereas the bootstrapped direct, indirect, and total effects summarize contributions across the corresponding direct and mediated pathways (Fig.4). The complete decomposition of the indirect effects of island area on EF and their bias-corrected and accelerated (BCa) confidence intervals are provided in Table S2. [318-322]

Section 3.2 (Results):

The corresponding combined indirect effect was −0.044 (95% BCa CI: −0.213–0.103), resulting in a total effect of 0.395 (95% BCa CI: 0.314–0.532; Fig. 4). Among the individual indirect pathways, the serial pathway from island area through observed woody plant diversity and structural complexity to EF was positive (0.045, 95% BCa CI: 0.010–0.130). In contrast, the pathway from island area through structural complexity alone was negative (−0.066, 95% BCa CI: −0.185–−0.014). The remaining indirect pathways were comparatively small, with confidence intervals overlapping zero. Estimates for all decomposed pathways are provided in Table S3. [366-372]

Section 4.2 (Discussion):

The direct path was associated with at high standardized effect size (Std.Beta = 0.44, p < 0.001) (Fig. 4)., indicating that the positive island area – EF relationship was predominantly represented by the direct island area – EF effect in this model. [426-427]

The positive serial pathway from island area through observed woody plant diversity and structural complexity was significant but comparatively small. Moreover, it was counteracted by other indirect pathways, including a negative pathway from island area through structural complexity to EF (Table S3). These results suggest that the positive island area – EF relationship in the SEM was dominated by the direct island area – EF, whereas the indirect pathway through observed woody plant diversity and structural complexity provides evidence for one possible, but limited mechanism contributing to this relationship. Further underlaying patterns and mechanisms deserve additional scientific attention. [463-469]

Section 5 (Conclusion):

Tree-island area enhances evaporative cooling through both direct and indirect pathways, primarily through a direct area – EF effect. An indirect serial pathway through observed woody plant diversity and structural complexity provides evidence for a secondary mechanism contributing to this relationship. [506-508]

Minor Comments

1. Spatial context: Figure 1 shows the distribution of the tree islands but does not include the plantation boundary, major roads, or surrounding land cover. Some additional spatial context would help readers assess whether proximity to plantation edges or contrasting land-cover types could covary with island-area class or influence the thermal measurements. If this information is available in the original experimental-design paper, a brief summary and reference here may be sufficient.

    We thank the reviewer for this helpful suggestion. The experiment was established within a large, approximately 140-ha oil-palm monoculture, and the experimental treatments were randomly distributed throughout the plantation, with a minimum distance of 85 m between islands (Teuscher et al., 2016). Thus, all experimental islands considered here were embedded within the same oil-palm plantation matrix rather than being positioned near contrasting land-cover types or plantation boundaries. At the spatial extent shown in Figure 1, the plantation boundary is sufficiently distant from the experimental islands that it does not provide meaningful spatial context for the thermal measurements, and including it would substantially expand the map extent without improving interpretation of the experimental comparisons. Other land-cover types, including rubber plantations, agricultural areas, and secondary forest, occur in the broader landscape, but these are located outside the immediate experimental matrix and were not expected to covary systematically with island-area class.

2. Temperature endpoints. Please clarify whether the DATTUTDUT hot and cold endpoints were calculated separately for each island or from the complete mission-level orthomosaic. If they were calculated separately for each island, it would be helpful to assess sensitivity to the strongly unequal pixel numbers among area classes.

We thank the reviewer for raising this point. The DATTUTUDUT temperature endpoints were calculated separately for each tree island. Specifically, each LST orthomosaic was first clipped to the individual island polygon, after which QWater calculated the hot and cold endpoints from the pixels within that island. The hot endpoint corresponds to the 100th percentile and the cold endpoint to the 0.5th percentile of the island-specific LST distribution.

We agree that this procedure results in different numbers of pixels contributing to the endpoint estimates among island-area classes. We had previously assessed the relevant upper and lower tails of the island-specific LST distributions using the same high-radiation observations as in the main analysis. For each island, we calculated the mean LST, the mean of the 1% hottest pixels, and the mean of the 0.5% coldest pixels, and tested for differences among island-area classes using Kruskal–Wallis tests (Supplementary Materials 3). None of these metrics differed significantly among island-area classes (mean LST: $p = 0.50$; 1% hottest pixels: $p = 0.48$; 0.5% coldest pixels: $p = 0.48$). These results provide no evidence that the upper and lower tails of the island-specific LST distributions systematically differed among island-area classes, despite the different numbers of pixels available within the islands.

We have clarified in the Methods that the DATTUTUDUT endpoints were calculated separately for each island and now refer to this sensitivity analysis in the Results/Supplementary Materials (Chapter 5).

We re – vised section 2.4 (Methods):

For each observation, the LST orthomosaic was first clipped to the polygon of each tree island, and QWaterModel was then run separately for each clipped tree island raster. Thus, the hot and cold temperature endpoints were calculated from the pixels within each individual tree island. [234-236]

Section 3 (Results):

To assess whether the unequal numbers of pixels among island-area classes affected the island-specific LST endpoints, we compared mean LST, the mean of the 1% hottest and the mean of the 0.5% coldest pixels among area classes. None differed significantly (Kruskal-Wallis tests: $p = 0.50$, $p = 0.48$, $p = 0.48$, respectively, Supplements Fig.S2 – S4). [332-335]

3. Sample composition. The random-partitions analysis uses N=48, whereas the SEM uses N=52. Please summarize clearly which plots were included in each analysis and the rationale for the different sample sizes.

   We thank the reviewer for pointing out the difference in sample size between the two analyses. The EFForTS-BEE experiment comprises 56 tree-island islands in total. Of these, 52 islands belong to the factorial experimental design described by Teuscher et al. (2016), comprising 13 islands in each of the four island-area classes (5 × 5, 10 × 10, 20 × 20, and 40 × 40 m). These 52 islands include one island per area class with a planted-tree diversity level of zero, i.e. islands in which no trees were planted but which received the experimental management treatment. The remaining four islands (P53–P56) are 10 × 10 m business-as-usual controls, which were neither planted nor subjected to the experimental management treatment.

   The SEM therefore included all 52 islands belonging to the experimental design, including the four diversity-level-0 islands. In contrast, the random-partition analysis included the 48 islands with planted trees (diversity levels 1, 2, 3, and 6), excluding both the four business-as-usual controls and the four experimental islands with a diversity level of zero. This corresponds to 12 planted islands per island-area class. The exclusion of the diversity-level-0 islands from the random-partitions analysis follows the treatment structure of that analysis, which was designed to evaluate the relationships among island area and the planted-tree communities within the random-partitions component of the experiment. The experimental design and distinction between the 52 experimental plots and the four business-as-usual controls are described in detail by Teuscher et al. (2016).

   We have clarified the sample composition and the rationale for the different sample sizes in the Methods.

   We re - vised section 2.9 (Methods):

   In EFForTS-BEE, the random partitions model focused on 48 plots, comprising 12 plots per island-area class. The four plots without planted trees and the four “business-as-usual” control plots were not part of the analysis, since they do not contribute to the species richness gradient. [296-297]

   The “business-as-usual” plots were not included in the SEM because their continued management might introduce confounding effects; the SEM therefore included the 52 plots belonging to the factorial experimental design, including the four plots without planted trees. [325-327]

4. Language and typographical corrections. Please replace the noun “analyzes” with “analyses” throughout, correct “Priestly–Taylor” to “Priestley–Tylor,” and correct “pievewiseSEM” to “piecewiseSEM.” The sentence “The direct path was associated with at high standardized effect size” should read “The direct path was associated with a high standardized effect size.”

We thank the reviewer for pointing out these language and typographical errors, we have re – vised the manuscript accordingly.