Dear Editor and Reviewer #1,

We sincerely thank the editor and the two anonymous reviewers for their careful evaluation of our manuscript and for providing constructive and insightful comments. We greatly appreciate the time and effort they devoted to reviewing our work. The comments were extremely helpful for improving the clarity, balance, and scientific focus of our manuscript.

In response to the reviewers' comments, we have substantially revised the manuscript. The major revisions include: (1) reframing the study as intermittent multi-year observations rather than continuous long-term measurements; (2) revising the title, objectives, Results, Discussion, and Summary and conclusions to emphasize methodological insights, uncertainties, and limitations; (3) clarifying the sampling schedule and treatment of missing data within the daytime sampling window; (4) revising the drone-related sections to describe the measurements as supplemental and exploratory rather than fully integrated with the study or used for validation; (5) expanding the MEGAN Methods section to clarify the model calculation period, activity factor settings, and their pairing with observations; and (6) revising the Supplementary Materials, including figure and table captions, to clarify that the statistics are based on available daytime samples under intermittent sampling conditions.

Below, we provide point-by-point responses to each comment. The reviewers' comments are reproduced in italics, followed by our responses. Changes in the revised manuscript are indicated in the marked-up version, and a clean revised manuscript and revised Supplementary Materials are also provided.

Response to Reviewer #1

We sincerely thank the reviewer for carefully reading the manuscript several times and for providing detailed and constructive comments. We appreciate the reviewer's concerns regarding the original framing of the study, especially the use of the term "long-term" and the emphasis on seasonal and interannual comparisons. We agree that our dataset, consisting of 34 intermittent sampling days from June 2023 to October 2025, should not be described as a continuous long-term flux dataset. We also agree that the drone measurements should not be presented as fully integrated with the tower observations because they were conducted only on selected days and for short sampling durations.

Following the reviewer's suggestion, we have substantially revised the manuscript to emphasize the methodological aspects of the study. Specifically, we now describe the dataset as "intermittent multi-year observations" rather than "long-term observations," and we present the drone measurements as "supplemental drone-based measurements" rather than "integrated tower and drone observations." We have also revised the Title, Abstract, Introduction, Methods, Results, Discussion, and Summary and conclusions to focus on the potential and limitations of tower-based multi-height BVOC observations, supplemental drone sampling, and MEGAN comparison under sampled daytime conditions.

We also reduced the emphasis on seasonal and interannual interpretations. Seasonal grouping is now used only to summarize model–observation differences under sampled phenological and meteorological conditions, and we explicitly state that the dataset should not be interpreted as a complete seasonal or interannual characterization of BVOC emissions at the site. In addition, we revised the figure captions and Supplementary Materials to clarify that monthly and daily statistics were calculated from available daytime samples collected during the intermittent sampling campaign.

Point-by-point responses to Reviewer #1

Comment 1: The introduction should be shorter and more concise.

Response: We thank the reviewer for this suggestion. We agree that the original Introduction was too lengthy and contained several parts that were not directly connected to the main objectives of the study. We have therefore substantially shortened and reorganized the Introduction. In the revised version, we focus more clearly on the need for canopy-scale BVOC observations in suburban forests, the methodological uncertainties associated with flux measurements, the potential use of supplemental drone-based sampling, and the motivation for comparisons with MEGAN.

Comment 2: The citations should be placed in chronological order.

Response: We thank the reviewer for pointing this out. We have checked the citation order throughout the manuscript and revised all citations so that they are placed in chronological order where appropriate.

Comment 3: Revise the methodological section as mentioned above.

Response: We agree with the reviewer that the methodological section needed substantial revision to clarify the observational design and the limitations of the measurements. In response, we revised the Methods section to describe the dataset as "intermittent multi-year observations" rather than "long-term observations." We also revised the description of the drone measurements from "integrated tower and drone observations" to "supplemental drone-based measurements." In addition, we clarified the sampling schedule, the number of sampling days, the treatment of missing samples within the daytime sampling window, the use of field blanks and recovery tests, the assumptions and uncertainties associated with the aerodynamic gradient method, the limitations of drone-derived flux estimates without direct turbulence measurements, and the details of the MEGAN calculations and activity factor settings.

Comment 4: Reduce the number of plots/images; so many are unnecessary and become distracting.

Response: We thank the reviewer for this helpful suggestion. We revised the figure structure to reduce the number of figures in the main manuscript and to keep the main text focused on the key results supporting the methodological interpretations of the study. Specifically, Figures 3, 6, and 12 in the original manuscript have been moved to the Supplementary Materials. Figure 13 in the original manuscript was removed because the revised manuscript no longer emphasizes interannual trends or the 2025 summer as evidence of long-term change. Following a suggestion from Reviewer #2, the original Figure S6 was moved to the main manuscript as Figure 9 because it directly supports the comparison between observed and MEGAN-calculated fluxes. We also revised the captions in both the main manuscript and the Supplementary Materials to clarify that the statistics are based on available daytime samples during the intermittent sampling campaign.

Comment 5: Rewrite the results and discussions with a methodological paper in mind.

Response: We agree with the reviewer's assessment that the main contribution of the manuscript lies in the methodological application and evaluation of tower-based multi-height measurements, supplemental drone-based sampling, and MEGAN comparison, rather than in providing a complete long-term seasonal or interannual flux record. Accordingly, we have substantially rewritten the Results and Discussion. In the revised manuscript, we avoid interpreting the dataset as a continuous long-term record and instead describe it as intermittent multi-year observations. We reduced the emphasis on seasonal and interannual comparisons and now use seasonal grouping only to summarize tendencies and model–observation differences under sampled conditions. The Discussion has been reorganized to focus on environmental controls under sampled daytime conditions, spatial representativeness and uncertainty of supplemental drone observations, and the possible causes of discrepancies between tower-based fluxes and MEGAN estimates. We also added a clearer discussion of the limitations and future improvements needed for tower- and drone-based BVOC flux observations.

Dear Editor and Reviewer #2,

We sincerely thank the editor and the two anonymous reviewers for their careful evaluation of our manuscript and for providing constructive and insightful comments. We greatly appreciate the time and effort they devoted to reviewing our work. The comments were extremely helpful for improving the clarity, balance, and scientific focus of our manuscript.

In response to the reviewers' comments, we have substantially revised the manuscript. The major revisions include: (1) reframing the study as intermittent multi-year observations rather than continuous long-term measurements; (2) revising the title, objectives, Results, Discussion, and Summary and conclusions to emphasize methodological insights, uncertainties, and limitations; (3) clarifying the sampling schedule and treatment of missing data within the daytime sampling window; (4) revising the drone-related sections to describe the measurements as supplemental and exploratory rather than fully integrated with the study or used for validation; (5) expanding the MEGAN Methods section to clarify the model calculation period, activity factor settings, and their pairing with observations; and (6) revising the Supplementary Materials, including figure and table captions, to clarify that the statistics are based on available daytime samples under intermittent sampling conditions.

Below, we provide point-by-point responses to each comment. The reviewers' comments are reproduced in italics, followed by our responses. Changes in the revised manuscript are indicated in the marked-up version, and a clean revised manuscript and revised Supplementary Materials are also provided.

Response to Reviewer #2

We sincerely thank the reviewer for carefully reading the manuscript and for providing detailed and constructive comments. We greatly appreciate the reviewer's efforts to identify unclear descriptions in the observational design, drone measurements, MEGAN calculations, and the interpretation of the results. These comments were extremely helpful in improving the clarity, balance, and scientific focus of the manuscript.

In response, we have substantially revised the manuscript. We clarified that the observations were intermittent multi-year measurements rather than continuous long-term measurements, specified the sampling times and gaps within the 10:00–16:00 JST sampling window, and revised the interpretation of monthly, seasonal, and interannual variability. We also revised the drone-related sections to emphasize that the drone measurements were supplemental and exploratory, especially because turbulence was not measured at the drone sampling locations, and tower-derived micrometeorological parameters were used for the drone-based flux calculations.

We also substantially revised the MEGAN section. We clarified the calculation period, time resolution, activity factor settings, and the pairing of model outputs with observations. We now state that the MEGAN calculations were conducted at a 10-min time resolution for the daytime period from 10:00 to 16:00 JST for each day from 1 June 2023 to 31 October 2025, and that model outputs corresponding to the actual sorbent tube sampling times were extracted and paired with the observed fluxes. We also clarified that the temperature history term T_{240} was calculated from continuous meteorological data for the preceding 240 h, so that antecedent thermal conditions were included in the temperature response.

We agree that the MEGAN comparison should not be interpreted as a complete model validation, especially given the intermittent sampling design and uncertainties in AGM-based observed fluxes. We therefore revised the manuscript to describe the comparison as a baseline model–observation evaluation under sampled daytime conditions. We also added a discussion of observational and model uncertainties, including concentration-gradient uncertainties, eddy diffusivity, stability corrections, canopy representativeness, basal emission factors, activity factors, and the canopy microclimate.

Regarding the use of *Q. serrata* specific basal emission rates, we agree that sensitivity tests using species specific and sun/shade-specific emission factors are important. However, such tests require additional assumptions regarding canopy structure, sunlit and shaded leaf fractions, leaf temperature, and scaling from leaf-level emission capacity to canopy-scale fluxes. We intend to collaborate with modeling experts on a separate project to conduct a full validation of such models; however, this falls outside the scope of the present paper. This paper focuses on methodological insights from tower-based multi-height observations, supplemental drone sampling, and a baseline MEGAN comparison. We have clarified this limitation in the revised discussion and identified species-specific MEGAN sensitivity tests as an important next step for future work.

Point-by-point responses to Reviewer #2

Comment 1

Lines 49-50: There are more recent global estimates than Guenther et al. (1995).

Response: We thank the reviewer for this comment. We have revised the text and replaced the older global estimate with the more recent estimate reported by Guenther et al. (2012).

Comment 2

Lines 79-90: It is important to mention the limitations of enclosure methods regarding scaling up to canopy emissions, but it is also important to mention the limitations of meteorological methods, which can be heavily influenced by complex canopies and terrains and, specifically for BVOCs, may not capture information on in-canopy fast chemical reactivities. Micrometeorological methods can be effective, but their limitations can be mitigated by combining measurements with enclosure methods.

Response: We thank the reviewer for this important suggestion. We have revised the Introduction to describe not only the limitations of enclosure methods in scaling leaf- or branch-level measurements to canopy-scale emissions, but also the limitations of micrometeorological methods. In particular, we now mention that micrometeorological flux estimates can be affected by complex canopy structure, terrain, roughness sublayer processes, and turbulent exchange assumptions. We also added that combining complementary approaches, such as enclosure measurements, vertical profiling, and spatially distributed sampling, can help to constrain these uncertainties.

Comment 3

Lines 148-151: In the introduction, you mention that this is a long-term study, but in M&M, you say that BVOC sampling was conducted for 1–3 days each month from June 2023 to October 2025, with some gaps due to equipment maintenance or power outages. This does not reduce the value and importance of more than two years of measurements, but I think one should be more careful with stating long-term measurements when only a few days per month (with some gaps) are considered. I suggest saying monthly measurements rather than just saying "long-term". I also want to mention that taking sorbent tube samples is really time-consuming, both in the field and in the lab, so doing this every month for more than two years is for sure a big effort.

Response: We thank the reviewer for this important comment. We agree that the original wording could have given the impression of continuous long-term measurements, which was not the case. We have therefore revised the manuscript throughout to describe the dataset as “intermittent multi-year observations” rather than continuous long-term observations. We also revised the Title, Abstract, Introduction, Methods, Results, Discussion, and Summary and conclusions to avoid overinterpreting the dataset as a complete seasonal or interannual record. We appreciate the reviewer’s recognition of the effort required for repeated sorbent tube sampling over more than 2 years.

Comment 4

Lines 164-165: Did you have a grid of LAI measurements? How were these measurements done? More details are needed here.

Response: We thank the reviewer for this comment. We have clarified that the LAI was measured at four fixed locations around the tower using a plant canopy analyzer, and that the average value from these locations was used. We also clarified that these LAI values represent local canopy conditions around the tower rather than those of the entire FM Tama forest.

Comment 5

Lines 171-175: Did you test for a potential loss of BVOCs due to the dehumidification tube and ozone scrubber cartridge?

Response: We thank the reviewer for this comment. We conducted preliminary laboratory recovery tests to evaluate possible losses of the target BVOCs associated with the dehumidification tube and ozone scrubber cartridge. We have added this information to the revised Methods section.

Comment 6

Lines 179-180: Why did the sampling time range from 10 to 30 min? Did you have a blank sorbent tube during sampling?

Response: We thank the reviewer for pointing out this unclear description. Tower-based sorbent tube sampling was generally conducted with 30 min integrated samples. During the supplemental drone-based observations, however, the sampling duration was shortened to 15 or 17 min to match the drone sampling duration and operational constraints. We have clarified this in the revised manuscript. We also clarified the blank handling in response to Comment 7.

Comment 7

Lines 196-198: Why was only the travel blank subtracted from the samples? I understand that contamination may occur during transport and storage, but a blank during the sampling would give

you more information about potential diffusion/contamination of your sampling system.

Response: We apologize for the confusion caused by the terminology in the original manuscript. The "travel blanks" described in the original text were actually field blanks. Field blanks were used to evaluate potential contamination during transport, storage, field handling, and exposure associated with the sampling procedure. For each sampling campaign, blank sorbent tubes were transported to the field together with the sample tubes. During sampling preparation and after sample collection, the field blank tubes were uncapped and recapped at the same time as the sample tubes. During active sampling, the field blank tubes were kept capped and placed near the active sampling tubes. The field blanks were subsequently transported, stored, and analyzed in the same manner as the actual samples. The blank values were subtracted from the measured values of the corresponding samples. We have corrected the terminology from "travel blank" to "field blank" and clarified the procedure in the revised manuscript. We also note that these field blanks were not connected to the active sampling line; therefore, possible contamination or losses within the inlet and sampling lines were not fully evaluated.

Comment 8

Lines 283-284: Is the 10-min time resolution for the AGM method?

Response: We thank the reviewer for this comment. The 10 min time resolution referred to the MEGAN calculations, not to the AGM sampling interval. In the revised manuscript, we clarified that MEGAN calculations were conducted at a 10 min time resolution for the daytime period from 10:00 to 16:00 JST for each day from June 1, 2023 to October 31, 2025. For comparison with the observations, model outputs corresponding to the actual sorbent tube sampling times were extracted and paired with the observed fluxes.

Comment 9

Line 297: I suggest moving the setting of activity factors related to environmental conditions used in MEGAN to the main text. This is important information about the model settings.

Response: We thank the reviewer for this helpful suggestion. In the revised manuscript, we moved the description of the MEGAN activity factor settings from the Supplement to the main text. The revised Methods section now describes the activity factors settings.

Comment 10

Lines 312-313: In the M&M, it states that the sampling period ranged from June 2023 to October

2025, which is less than 3 years. In addition, according to Table S1, some months had only one day of measurements. It is important to be very clear about the number of measurements. The way it is written may lead the reader to think of 3 years of continuous measurements, which is not the case, as I understood it. Also, Table S2 should include the number of samples/measurements for each month. What is the variation of the means? You should include the standard deviation. I did not see it in Figure 3, nor in Table S2.

Response: We agree with the reviewer's concern. We revised the manuscript throughout to avoid giving the impression of continuous three-year measurements. We now clearly state that BVOC samples were collected on 34 sampling days during the intermittent multi-year campaign from June 2023 to October 2025. We also revised the figure and table captions to clarify that monthly means were calculated from available daytime samples collected on the observation days. Following the reviewer's suggestion, we added standard deviations to Table S3 summarizing monthly BVOC volume mixing ratios. To reduce the number of figures in the main text, as suggested by Reviewer #1, some of the figures have been moved to the Supplement.

Comment 11

Lines 333-334: This is probably more appropriate for the discussion section.

Response: We thank the reviewer for this suggestion. The sentence was intended to summarize the observational result that monoterpenes volume mixing ratio were much lower than those of isoprene and showed unclear vertical gradients. We retained a shortened version in the Results because it provides the basis for focusing the subsequent flux analysis and Discussion on isoprene. We have revised the wording to keep the statement descriptive and avoid overinterpretation.

Comment 12

Line 343: Again, the sampling period in Figure 4 is stated as 3 years, which seems incorrect. Why did March show lower mixing ratios than December, January, and February? Were all the data from 10 am to 4 pm combined for the plots in Figure 4? Is the high variability (especially in the summer) due to daily temperature variability? How did you treat the variation in data points when comparing seasons? For example, in the summer your N is 112, while in spring this is 36. How can you distinguish daily variation from seasonal variation with such an unbalanced sample size across the sampling period?

Response: We thank the reviewer for this important comment. We have revised the text and figure captions to avoid describing the dataset as a continuous 3-year record. We clarified that the boxplots were based on available daytime samples collected between 10:00 and 16:00 JST on the observation days. We also revised the interpretation of monthly and seasonal differences. As the number of sampling days and data points differed among months and seasons, we no longer interpret these differences as representing complete seasonal or interannual variability. Instead, we describe them as tendencies observed under the sampled daytime conditions. We also added text explaining that the dataset should not be regarded as a complete climatology of BVOC emissions at the site.

Comment 13

Line 381: Here, you say two and a half years. Overall, the sampling period lacks consistency throughout the text.

Response: We thank the reviewer for pointing this out. We have revised the manuscript throughout to ensure consistency. The study period is now described as an intermittent multi-year observation campaign from June 2023 to October 2025, with a total of 34 sampling days.

Comment 14

Line 389: In Figure 5, it is possible to see that for some months, there is only one measurement. For example, for June, you have data from June 29 (2023), June 11 (2024), and June 17 (2025), so the June plot in Figure 4 has only three data points. How may the meteorological conditions of these three days of different years have affected the samples you got? From Table 1, it is evident that the meteorological conditions differed, particularly in temperature. Also, what is the variability (at least within the day) in fluxes?

Response: We agree with the reviewer. As some months included only a few sampling days, meteorological differences among observation days can strongly affect the monthly summaries. We therefore have revised the manuscript to avoid interpreting monthly differences as representative of monthly climatology or robust interannual variability. We also clarified that daily averages and standard deviations in Table 1 were calculated from available samples within the daytime sampling window, so they represent within-day variability under the sampled conditions. We added text in the main body explaining that monthly, seasonal, and annual features should be interpreted as tendencies observed under the sampled conditions rather than as complete climatological patterns.

Comment 15

Lines 420-433: more appropriate to the discussion section.

Response: We thank the reviewer for this suggestion. We have moved most of the explanatory physiological discussion on temperature, light, and isoprene emission mechanisms to the revised

Discussion section.

Comment 16

Lines 441-442: Do you have a citation for this statement? Also, what time did you sample for the horizontal comparison? Can the variation be explained by combining samples from different daytimes?

Response: We thank the reviewer for this comment. This statement was based on the "Manual for Measurement of Hazardous Air Pollutants" issued by the Ministry of the Environment, Government of Japan, which describes quality control procedures for hazardous air pollutant measurements in Japan. We have revised the Methods section to clarify that the tower and drone measurements used for the horizontal comparison were conducted simultaneously at the same height. We have also revised the interpretation of the drone observations. As the drone measurements were limited to selected days and short sampling periods, we no longer describe them as integrated tower–drone observations. Instead, we describe them as supplemental measurements used to examine short-range horizontal variability around the tower.

Comment 17

Lines 463-465: Why do you have three ratios (0.72, 0.98, and 0.85) for only two distances (15 and 30 m)? I did not get that.

Response: We apologize for the unclear description. We revised the text to clarify that the average tower/drone flux ratios were 0.72, 0.98, and 0.85 for the 15 m dataset, the 30 m dataset, and the all-distance dataset, respectively. The all-distance dataset represents the combined 15 and 30 m drone sampling locations.

Comment 18

Line 474: Did the parameters to calculate flux with the AGM vary on the tower when comparing heights 23 and 30? Did you assume that the parameters to calculate flux with the AGM from drone samples were fixed and the same as 30m at the tower? If so, this could explain the low variability between drone and tower measurements. If I understand correctly, turbulence is assumed to be the same between tower and drone measurements; in this case, the variability between tower and drone fluxes is driven solely by variability in mixing ratios. How a potential variation in turbulence between tower and drone point samples could influence your flux calculations. I understand that you do not have turbulence data for the drone sample points, but it is important to at least consider

potential differences and elaborate on that. You briefly mention it in lines 488-490, but more is needed to explain the caveats of this method.

Response: We thank the reviewer for this important comment. We agree that the drone-derived flux estimates have substantial uncertainty because turbulence was not measured at the drone sampling locations. In the revised manuscript, we have clarified that micrometeorological parameters measured at the tower were used for the drone-based AGM calculations because the drone could not carry a three-dimensional ultrasonic anemometer due to payload limitations. Thus, the drone-based calculation assumes similar micrometeorological conditions between the tower and drone sampling locations. Under this assumption, differences between tower- and drone-derived flux estimates mainly reflect differences in concentration gradients and sampling height intervals, while possible horizontal variability in turbulence cannot be quantified. We have revised the Methods, Results, Discussion, and Summary and conclusion sections to state that the drone-derived fluxes should be interpreted as exploratory estimates of spatial representativeness and methodological uncertainty under tower-derived turbulence conditions, rather than as an independent quantitative validation of tower-based fluxes.

Comment 19

Lines 491-497: This method depended on tower measurements of turbulence. I suggest being more careful about indicating the method for increasing the spatial representativeness of flux observations without turbulence measurements on the drone.

Response: We agree with the reviewer. We have revised the manuscript to avoid suggesting that drone-based flux estimates can independently improve the spatial representativeness of flux observations without turbulence measurements. We now clearly state that the drone-derived flux estimates depend on tower-derived micrometeorological parameters and should be treated as exploratory. The revised text emphasizes that future drone-based BVOC flux studies would benefit from simultaneous multi-height sampling and direct turbulence measurements on or near the drone platforms.

Comment 20

Lines 510-512: Did you run the model continuously and then extract the results from 10 am to 4 pm? Or did you run the model only for these hours? This is not clear. By not continuously accounting for weather and LAI conditions, you do not consider the effects of previous days.

Response: We thank the reviewer for pointing out this unclear description. We have revised Section 2.7 to clarify how the MEGAN calculations were conducted and how the model outputs were paired with observations. In the revised manuscript, we state that the MEGAN calculations were conducted at a 10-min time resolution for the daytime period from 10:00 to 16:00 JST for each day from June 1, 2023 to October 31, 2025. The model outputs corresponding to the actual sorbent tube sampling times were then extracted and paired with the observed fluxes. We also clarified that the temperature history term T_{240} used in γ_T was calculated from continuous meteorological data for the preceding 240 h, rather than from daytime data only. Thus, antecedent thermal conditions before each daytime model time step were considered. We further state that the comparison represents model–observation agreement during sampled daytime periods and should not be interpreted as an evaluation of full-day emission estimates.

Comment 21

Lines 521-526: It is not a surprise that F_{cal}, temperature, and PPFD have good correlation, because MEGAN was mostly driven by temperature and PPFD variations.

Response: We thank the reviewer for this comment. We agree and have revised the text to avoid overstating this result. In the revised manuscript, we state that the strong relationships between F_{cal} and γ_T or γ_P primarily reflect the structure of the MEGAN algorithm, in which temperature and light activity factors are the primary drivers of daytime isoprene emission estimates.

Comment 22

Lines 534: Here again, the way weather conditions were treated to run MEGAN makes a difference. If MEGAN simulations were continuous (i.e., with hourly data), they could better capture the effects of previous days than selecting only the weather conditions from 10 am to 4 pm. This needs to be clarified in the text. Also, the small number of observations can limit comparability between observations and model estimates.

Response: We thank the reviewer for this important comment. As described in response to Comment 20, we revised Section 2.7 to clarify the MEGAN calculation and its pairing with observations. We also clarified that T_{240} was calculated using continuous meteorological data for the preceding 240 h, so that antecedent thermal conditions were included in γ_T. In addition, we revised the Results and Discussion sections to state that the comparison is limited by the small number of observations and should be interpreted as a baseline model–observation comparison under sampled daytime conditions, rather than as a complete seasonal or annual model validation.

Comment 23

Line 537: If you consider data from May to October for the analysis in Table 2, you have less data

in the Spring and Autumn than in the Summer. How did you treat that in your analysis?

Response: We thank the reviewer for this important comment. We agree that the number of paired data points differed among seasons, with fewer data in spring and autumn than in summer. In the revised manuscript, we now explicitly show the number of paired data points used for each seasonal comparison in Table 2. We did not apply weighting or resampling to equalize the number of data points among seasons. Instead, the seasonal statistics were calculated separately within each sampled season and are used only to describe model–observation differences under the sampled conditions. We clarified that these statistics should not be interpreted as a complete seasonal evaluation of MEGAN performance.

Comment 24

Lines 545-554: Is this analysis impacted by the high variability of observation data, especially in the 2025 summer? Do you have all the data between 10 am and 4 pm for all measurement days? Could a potential gap within the day increase this variability?

Response: We thank the reviewer for this important comment. We agree that the high variability of the observed fluxes, especially during the summer 2025 sampling days, may have affected the model–observation comparison by increasing RMSE and scatter. We also clarified that complete 10:00–16:00 observations were not available for all sampling days. As described in the revised Section 2.2 and Table S1, sampling times and durations differed among observation days, and when samples were missing within the daytime sampling window, daily averages were calculated only from available paired observations. For the MEGAN comparison, we used only time-synchronized paired data for which both F_{obs} and F_{cal} were available at the corresponding sampling times. We have also added text explaining that gaps within the daytime sampling window may increase uncertainty in daily statistics because short-term variability within the day was not always fully captured. Therefore, the comparison should be interpreted as model–observation agreement under sampled daytime conditions, rather than as a complete evaluation of diurnal or seasonal model performance.

Comment 25

Lines 560-561: What do you mean by low observation conditions?

Response: We apologize for the unclear wording. We have revised the text to avoid the ambiguous phrase “low observation conditions.” The revised sentence explains that the positive intercepts in the seasonal regressions suggest that MEGAN tended to estimate non-negligible fluxes even when the observed fluxes were low.

Comment 26

Lines 597-600: This, together with Figure 13, should be in the results section. In Figure 13, how do you explain the low fluxes at high temperatures for the year 2023? In lines 607-609, you mention that in 2025, the rainy season ended unusually early; this should be supported by data and discussed in greater detail alongside the results.

Response: We thank the reviewer for this comment. After reconsidering the limited number of sampling days, we agree that the dataset is not sufficient to support a detailed interannual comparison or a robust interpretation of the 2025 summer as evidence of a long-term trend. We therefore have removed the original Figure 13 and the associated interpretation. The revised manuscript no longer claims that the high 2025 fluxes demonstrate an interannual trend. Instead, we state more cautiously that several high-flux events were observed under hot summer sampling conditions and that further observations, including more detailed vegetation surveys, footprint analysis, and additional sampling campaigns, are needed to explain the low fluxes observed under some high-temperature conditions.

Comment 27

Lines 609-611: I do not think you have sufficient data to claim that you can detect a significant upward trend in isoprene emissions under the climatic conditions in Japan in 2025, nor to make interannual comparisons based on long-term observations, because you do not have enough data for such analyses.

Response: We agree with the reviewer. We have removed the statement claiming the detection of a significant upward trend in isoprene emissions in 2025 and revised the manuscript to avoid making interannual conclusions based on the intermittent dataset. The revised text states only that high fluxes were observed on several hot summer sampling days in 2025, while emphasizing that the dataset cannot separate interannual variability from day-to-day meteorological variability.

Comment 28

Lines 632-639: This is part of the results and not the discussion.

Response: We thank the reviewer for this suggestion. We have moved this statement from the Discussion to the Results section and revised the wording to keep it descriptive. The revised Results section now states that temperature and light were important drivers of F_{obs} under the sampled daytime conditions, but that they did not fully explain the observed flux variability.

Comment 29

Lines 654-658: This is disconnected from the previous and following paragraphs.

Response: We thank the reviewer for pointing this out. We removed the disconnected paragraph and reorganized Section 4.3 to improve the flow of the Discussion. The revised section now discusses the possible causes of model–observation discrepancies in a more coherent order: AGM-based observational uncertainty, basal emission factor uncertainty, temperature response uncertainty, and canopy-scale representativeness.

Comment 30

Lines 683-728: The conclusion section is a repetition of the results and not a conclusion.

Response: We thank the reviewer for this comment. We have rewritten the section as "Summary and conclusion." The revised version summarizes the main findings more concisely, emphasizes the limitations of the intermittent sampling design, and highlights the methodological implications and future research needs rather than repeating the Results in detail.