

General answers to the manuscript entitled: Divergent mercury sequestration dynamics in tropical dry and moist broadleaf forests

We thank the reviewers for their valuable comments, the time they dedicated to evaluating our manuscript, and their constructive suggestions, which will help us substantially improve the manuscript. Overall, the main comments concerned the use of terminology that more accurately reflects the measurements performed in this study and the need to distinguish more clearly between measured variables and inferred processes. Accordingly, terms such as “soil Hg stocks” and “litter-derived Hg inputs to soil,” which were not directly quantified with our sampling design, will be revised throughout the manuscript. We will also moderate mechanistic and causal interpretations where they cannot be fully supported by our data, and explicitly acknowledge the limitations of the study where appropriate. Our point-by-point responses are provided below.

Reviewer 1

Summary Assessment

This study provides valuable baseline data on foliar Hg concentrations in two contrasting tropical forest types (TDBF and TMBF). The authors report clear seasonal patterns in TDBF and their absence in TMBF, and they discuss possible mechanisms. However, several critical weaknesses limit the strength of the conclusions. These include: (1) a mismatch between measured traits (morphological) and the physiological processes believed to control Hg uptake; (2) untested mechanistic hypotheses presented as plausible explanations without supporting data; (3) insufficient phenological data to support claims about leaf age; (4) lack of replication for the deforested/reforested site; and (5) absence of multivariate analysis for soil–litter Hg relationships.

Below are specific recommendations. These are the ones I consider most important, and I ask the authors to pay attention to them.

1. Mismatch between measured traits and physiological drivers of Hg uptake

Section 3.2.3 - *"Specific leaf area index (SLA) and stomatal density index (SDN) were weakly correlated with foliar Hg concentration... high intra-individual variability ... likely weakens these relationships. ... Teixeira et al. emphasized physiological processes such as net photosynthesis – rather than SLA – as key drivers."*

You measured SLA and SDN, found weak correlations, and cite Teixeira et al. (2018) to argue that physiological processes (e.g., net photosynthesis, stomatal conductance) are more important. However, did you actually measure stomatal conductance or photosynthesis on the same leaves/trees/seasons? If not, why include SLA and SDN at all? Presenting weak correlations as a finding risks misleading readers into thinking these traits were expected to be relevant. The current literature (including the mechanistic model you cite elsewhere) clearly shows that stomatal and mesophyll conductance – not static morphological traits – govern GEM dry deposition.

If physiological data are unavailable, remove the emphasis on SLA/SDN correlations or present them explicitly as a negative result that confirms the irrelevance of static morphological traits.

Alternatively, clearly state in the Discussion that SLA and SDN were measured because no study has yet analyzed SDN on the same leaf samples used for Hg analyses, and that the weak correlation underscores the need for physiological measurements in future work.

Our response: *Thank you for your valuable comment. We agree that presenting weak correlations might mislead readers. Therefore, we will remove the section “3.2.3 Foliar functional traits and their relationship with Hg accumulation in tropical forests” and state that the weak correlation found between SLA and SDN vs Hg in foliage underscores the need for physiological measurements in future work.*

2. Untested mechanisms for lower Hg in TDBF evergreens during dry season

Section 3.2.1 - *Three possible mechanisms proposed: (a) some "evergreen" trees are actually semi-deciduous (Borchert, 1994) – new leaves in dry season; (b) photochemical re-emission of Hg⁰ (up to 45%); (c) phloem translocation (5–15%) to woody tissues. Conclusion: "the complex water-dependent dynamics.....- ... further complicate identifying the main mechanism."*

You present three competing mechanisms but test none with your own data. The Borchert (1994) reference is useful, but do you have individual-level phenological data (e.g., percent canopy emptiness, leaf age class for each sampled tree) for at least some species? Without such data, the "new leaves in dry season" hypothesis is speculation. Regarding photochemical re-emission: did you measure irradiance and temperature at the canopy level? If not, you cannot estimate its potential contribution.

Present any available phenological data (even qualitative) for the evergreen species sampled.

If no such data exist, reframe the text from "the following mechanisms may contribute" to "the following mechanisms have been proposed in other studies, but none could be tested here."

Explicitly state the absence of irradiance/temperature data as a limitation for evaluating the re-emission hypothesis.

Our response: *During our second sampling campaign in the dry season, we observed that some evergreen trees that had full canopy cover during the first dry-season sampling campaign had fewer leaves and, in some cases, had lost all their leaves. As these changes were only documented during the second dry-season sampling campaign, we cannot provide comprehensive quantitative data on canopy changes between the two dry-season campaigns. Nevertheless, we will include in the Supplementary Material the available field observations of canopy cover and visually estimated leaf age (young or mature) recorded during our third sampling campaign. These observations also show which evergreen trees could be sampled during the first dry-season campaign but could not be sampled during the second because they had lost their leaves. During the wet-season sampling, evergreen trees generally had full canopy cover. These observations support the possibility that some leaves collected from evergreen trees during the dry season were newly flushed. However, because leaf age was not quantitatively determined, we will present this explanation as a hypothesis rather than a demonstrated mechanism. Regarding irradiance and temperature data, we will compile*

available site-specific or nearby meteorological data corresponding to our sampling periods. If suitable data are available, we will include them to provide additional environmental context for evaluating the photochemical Hg re-emission hypothesis. If such data are not sufficiently representative of canopy conditions at our sampling sites, we will explicitly acknowledge this limitation in the revised manuscript, as requested.

3. Atmospheric Hg concentration difference (1.2 vs. 0.6 ng m⁻³) – causal attribution Section 3.1 - “TDBF has twice the atmospheric Hg⁰ concentration of TMBF. Explanation: TMBF has denser canopy + higher humidity → higher stomatal conductance → higher Hg⁰ uptake → lower ambient Hg⁰. The manuscript states: “higher Hg⁰ was found in TDBF in both seasons.”

Your explanation is internally consistent: if TMBF captures more Hg⁰, local air becomes more depleted. However, you provide no evidence that the difference is driven by vegetation uptake rather than external sources (e.g., biomass burning, exposed soils, long-range transport). Without upwind measurements or atmospheric transport modeling (e.g., GEOS-Chem), the causal statement is premature.

Moderate the claim to: “the lower atmospheric Hg⁰ concentration in TMBF is consistent with higher dry deposition, but alternative explanations (e.g., differences in regional air mass history) cannot be ruled out.”

If possible, compare your values with nearby background sites or model simulations for the region.

Our response: Thank you for your valuable comment. We agree, therefore, we will moderate the claim to “The lower atmospheric Hg⁰ concentration in TMBF is consistent with higher wet deposition, but alternative explanations (e.g., differences in regional air mass history, canopy density in each forest) cannot be ruled out”, as requested.

4. Leaf age vs. seasonal effects (deciduous trees in TDBF: 4.7× lower Hg in dry season)

“Deciduous trees lose leaves in dry season; dry-season samples are newly flushed leaves → less accumulated Hg. Attributed to leaf age (phenology), not a seasonal difference in uptake rate.”

Your reasoning is correct and consistent with literature. However, you did not normalize Hg concentration by exposure time or leaf area. Without such normalization, the ambiguity remains: dry-season leaves could have lower Hg because (a) they are younger or (b) the uptake rate is lower in dry season (e.g., due to stomatal closure). To distinguish these, you would need leaves of the same chronological age sampled in different seasons, or a bud-marking experiment.

Acknowledge this as a limitation in the Discussion.

Suggest: “Future studies should normalize foliar Hg by time-integrated stomatal conductance or by leaf age to separate seasonal uptake dynamics from age-related accumulation.”

Our response: We agree, therefore, we will explicitly acknowledge this limitation in the discussion with the statement: “Future studies should normalize foliar Hg by time-integrated

stomatal conductance or by leaf age to separate seasonal uptake dynamics from age-related accumulation.", as requested.

5. Deforestation/reforestation site – lack of spatial replication The deforested and reforested site (30 years) is unique.

Without spatial replication, any claim about a "persistent effect of deforestation" is at best a **case study**. The manuscript must explicitly acknowledge this limitation. Language implying causality (e.g., "deforestation caused a 2.6× reduction") should be replaced with descriptive phrasing (e.g., "the reforested site showed 2.6× lower soil Hg, consistent with other studies of post-agricultural regeneration").

Add a sentence to the Limitations section: "The deforested/reforested site was not spatially replicated; therefore, our findings from this site should be interpreted as a case study rather than a definitive demonstration of deforestation effects."

Compare your results with published values from other post-agricultural regeneration sites in the tropics.

Our response: *We agree, we will therefore carefully revise any language implying causality and replace it with descriptive wording. In addition, we will explicitly acknowledge this limitation in the discussion and compare our results with available studies from other post-agricultural or regenerating tropical forest sites.*

6. Definition of "deciduous" in TMBF (no pronounced dry season)

"In TMBF, no seasonal variation in Hg. Deciduous trees had mature leaves in both seasons. Citation: Kearsley et al. (2024) – phenology regulated by drought and light. No information is provided on the duration/intensity of the dry season in TMBF, nor quantitative phenological data. "

What is the length and intensity of the dry season in your TMBF site (e.g., months with precipitation < 50 mm)? If there is no pronounced water stress, what triggers deciduousness? Without quantitative phenological data (e.g., percent canopy emptiness, leaf age distribution per species), the statement that "sampled leaves were mature in both seasons" is an assumption. This is a significant limitation because the very definition of "deciduous" loses meaning without a clear environmental trigger.

Provide a table with species-level phenological data (e.g., timing of leaf flush, duration of leafless period, if any).

If such data do not exist, state: "We assumed that leaves were mature based on typical phenology for these species, but we did not quantitatively assess leaf age. This is a limitation."

Our response: *We agree, as leaf age and species-level phenology were not quantitatively assessed, we cannot confirm the chronological age of the sampled leaves. We will therefore revise the corresponding statement, explicitly acknowledge this limitation, and clarify that leaf maturity was inferred from field observations rather than quantitatively determined. We will also highlight that future studies should include quantitative assessments of leaf age and species-level phenology to better resolve seasonal patterns in foliar Hg.*

7. Soil–litter Hg correlation – risk of spurious correlation and need for pathway analysis

"Consistent with previous studies identifying litter as the primary Hg source to forest soils ... our results showed a positive correlation between litter Hg inputs and soil Hg concentrations (0–30 cm depth-averaged) in both tropical forests (Fig. 6a)."

You present a positive correlation between litter Hg inputs and soil Hg concentrations. However, as you acknowledge earlier in the litter section, litter mass and Hg content are influenced by decomposition dynamics, leaching, re-emission, and microbial activity. A simple bivariate correlation does not account for these processes. Moreover, both litter Hg inputs and soil Hg could covary with a third factor – e.g., soil organic matter content, clay percentage, pH, or historical fire regime – without a direct causal link from litter to soil.

Critically, you also report that in TDBF, litter Hg concentration is 34% lower than mature living leaves in the wet season but 64% higher than living leaves in the dry season. This implies that litter does not simply reflect foliar Hg at abscission; post-senescence changes (leaching, decomposition, re-emission) are substantial. Therefore, correlating litter Hg *inputs* (mass × concentration) with soil Hg while ignoring these transformations may overestimate the direct contribution.

Perform a multiple regression or path analysis with soil Hg as the dependent variable and litter Hg input, soil organic matter, clay content, pH, and possibly decomposition rate proxies (e.g., C:N ratio) as independent variables.

If such analysis is not possible, add an explicit limitation: "Our correlation does not demonstrate causality; litter Hg inputs and soil Hg may be independently influenced by common environmental factors (e.g., organic matter accumulation). Moreover, post-senescence Hg losses from litter before incorporation into mineral soil remain unquantified."

Our response: *Thank you for your valuable comment, we will compile the available data collected during the study and assess whether they are sufficient to perform a multivariate analysis that accounts for potential covariates influencing the relationship between litter Hg standing stock and soil Hg concentrations. If the available data allow a robust analysis, we will include the results in the revised manuscript. If not, we will explicitly acknowledge that the observed correlation does not demonstrate causality and that litter Hg standing stock and soil Hg concentrations may be influenced by additional environmental factors. We will also clarify that post-senescence Hg losses from litter before incorporation into the mineral soil remain unquantified.*

8. Soil Hg stocks in deforested/reforested site – loss pathways unspecified

Soil Hg stocks (0–30 cm) are 2.6× lower in the deforested/reforested site after 30 years of regrowth.

Compared to other post-agricultural studies, is this loss typical or extreme? Which depth increment is most affected (e.g., 0–10 cm vs. 10–30 cm)? The possible loss pathways are (a) volatilization, (b) leaching, (c) erosion. Did you measure soil Hg⁰ flux (dynamic chamber) or Hg in drainage waters? If not, the discussion of pathways remains speculative.

Present depth-resolved data if available.

If no flux or leaching data exist, state: "We did not measure Hg volatilization or leaching; therefore, the relative importance of different loss pathways cannot be determined from our data."

Provide a theoretical mass balance or cite regional studies that have quantified these pathways.

Our response: *We agree, we will include an explanation of the data on the lower Hg concentration in soils from reforested areas at each depth. Since we did not measure Hg⁰ flux (dynamic chamber) or Hg in drainage waters, we will add the statement you mentioned above, noting that a limitation of our justification is that we did not measure volatilisation or leaching. In addition, we will compare our results with available studies from post-agricultural or regenerating tropical forest sites to provide context for the magnitude of the observed differences, as requested.*

9. Litter Hg accumulation seasonality – potential double-counting or misinterpretation of "inputs"

"In TDBF, litter Hg inputs were 4.7 times higher in the dry season ($31 \pm 6 \mu\text{g Hg m}^{-2}$) than in the wet season ($7 \pm 1 \mu\text{g Hg m}^{-2}$), primarily due to greater litter mass accumulation."

"We quantified forest-floor litter present at the time of sampling (end of each season), rather than trap-based litter. Our measurements, therefore, represent net litter accumulation, shaped not only by leaf shedding but also by decomposition dynamics."

You call these "litter Hg inputs," but what you measured is standing stock of Hg in litter at the end of each season, not cumulative inputs over time. Because decomposition is ongoing, a high standing stock in the dry season could result from (a) high input, (b) slow decomposition, or (c) both. You acknowledge this in the text, yet you repeatedly use the term "inputs" (e.g., "litter Hg inputs," "litter-derived Hg inputs"), which implies flux. This is a terminological inconsistency that could mislead readers.

Furthermore, in TDBF, the dry-season standing stock is 4.7× higher than wetseason standing stock. If decomposition is severely reduced in the dry season (as you suggest: "limited soil moisture likely constrains microbial activity, slowing decomposition"), then the difference may be driven more by reduced output than by increased input. You cannot separate these without litter trap data or decomposition experiments.

Replace "litter Hg inputs" with "litter Hg standing stock" or "litter Hg pool" throughout the manuscript when referring to forest-floor measurements. Reserve "inputs" for trap-based studies or explicitly define it as "nets accumulation (inputs minus decomposition losses)."

Add a sentence: "Our design does not allow partitioning of the dry-season litter Hg pool into contributions from increased leaf shedding versus decreased decomposition; both likely contribute."

Our response: *We agree, therefore, we will replace "litter Hg inputs" with "litter Hg standing stock" throughout the manuscript where appropriate. We will also clarify as a limitation that our experimental design does not allow us to distinguish whether the larger dry-season litter Hg standing stock resulted mainly from increased leaf shedding, reduced decomposition, or a combination of both processes, as requested.*

Minor Comments

10. Inconsistent reporting of sample sizes

TDBF litter: "dry-season litter was 3.6 times higher ($0.7 \pm 0.1 \text{ kg m}^{-2}$, $n = 24$) than wetseason litter ($0.2 \pm 0.0 \text{ kg m}^{-2}$, $n = 24$)."

"When comparing both forest types across seasons, the TMBF showed 1.3 times higher litter accumulation ($0.57 \pm 0.0 \text{ kg m}^{-2}$, $n = 48$) than the TDBF ($0.45 \pm 0.1 \text{ kg m}^{-2}$, $n = 6$)."

The sample size for TDBF across seasons drops from $n = 24$ per season (total 48) to $n = 6$ for the forest-type comparison. This is confusing. Are the $n=6$ representing plotlevel averages (6 plots H1–H6)? If so, state clearly: " $n = 6$ plots, with seasonal replicates already averaged." The current presentation may lead readers to think you lost 42 samples.

Harmonize sample size reporting. Specify in the figure legend or text whether n refers to individual litter samples or plot-level means.

Our response: *Thank you for your carefully reading. We will clearly specify throughout the manuscript when we refer to plot-level or individual litter samples to avoid any confusion.*

11. Comparison with global litter estimates – missing uncertainty

"TDBF litter Hg concentration is lower than the mean global litter estimate from various forests of $54 \mu\text{g kg}^{-1}$ (Wang et al., 2016), but similar to the mean estimate for deciduous needleleaf forests reported by Xu et al. (2022) of $36 \mu\text{g kg}^{-1}$."

The global mean ($54 \mu\text{g kg}^{-1}$) likely includes temperate and boreal forests with different species and atmospheric Hg concentrations. A direct comparison without accounting for these differences is of limited value. More informative would be a comparison with other tropical dry forests – but you correctly note no data exist. Consider rephrasing: *"Direct comparison with global means is complicated by differences in atmospheric Hg^0 , climate, and species composition; however, our TDBF values ($\sim 33\text{--}43 \mu\text{g kg}^{-1}$) fall below the global forest average, highlighting the need for more TDBF-specific data."*

Moderate the comparison as suggested above.

Our response: *Thank you for your valuable comment. As there is currently a lack of data on Hg concentration in litter, as it is stated in the manuscript, we will moderate the comparison, as you suggested, emphasising that direct comparison is complicated due to Hg^0 differences, climate, species, etc.*

12. Litter Hg vs. living leaf Hg – interesting but unquantified mass balance

"In TDBF, the average foliar Hg concentration of mature living leaves during the wet season ($50 \mu\text{g kg}^{-1}$) was 34% higher than in litter ... Conversely, in the dry season, Hg concentration in litter was 64% higher than in living leaves ($15 \mu\text{g kg}^{-1}$)."

This is a fascinating pattern. The dry-season litter being higher in Hg concentration than living leaves suggests that litter consists of older, senesced leaves that accumulated Hg over a

longer period (or that Hg is retained during initial senescence before later losses). The wet-season litter being lower than living leaves suggests dilution by young leaves or post-senescence leaching. This could be a key insight, but it is presented descriptively.

Consider adding a conceptual figure or a short hypothesis paragraph explaining the seasonal shift in the litter-to-living-leaf Hg ratio. For example:

- Dry season: Litter = older, Hg-enriched leaves shed early; living leaves = newly flushed, low Hg.
- Wet season: Litter = leaves that have already lost Hg via leaching/decomposition; living leaves = mature, Hg-rich.

Our response: Thank you for your valuable suggestion. We will further clarify and expand the discussion of this described pattern to better explain the observed seasonal differences in Hg concentrations between living leaves and litter. We will also consider including a conceptual figure to illustrate the proposed mechanisms, as suggested.

13. "No significant differences" – report exact p-values or effect sizes

"Hg concentrations in litter from the TDBF showed [...] Differences were not statistically significant at the plot level, except for H6 (Fig. S4 b, SM). The same pattern of a lack of significant seasonality effect in most plots (except G4) was found for TMBF."

Stating "not statistically significant" without p-values or effect sizes (e.g., Cohen's d) is insufficient. Given your modest sample sizes (n = 24 per season per forest? Or per plot?), lack of significance could reflect low power rather than true absence of effect.

Report p-values (e.g., $p = 0.12$, paired t-test) or provide confidence intervals for the seasonal differences. If power analysis was performed, state it. Otherwise, add a caveat: "Small sample sizes per plot limit our ability to detect modest seasonal effects."

Our response: Thank you for your comment. As requested, we will specify p-values for the correlations, which will improve our statements of the statistical significances found throughout the manuscript. In addition, we will acknowledge that the relatively small sample size within individual plots may limit our ability to detect modest seasonal effects.

14. TDBF litter Hg accumulation – potential overestimate due to dry-season slow decomposition

"Greater litter-derived Hg accumulation during the dry season may be linked to reduced decomposition rates, as dry conditions are known to slow litter breakdown."

This is an important warning, but it undermines the earlier statement that litter is the "primary Hg source to forest soils." If a large fraction of dry-season litter Hg is still present as *undecomposed litter* at the time of sampling, it may not have been transferred to the mineral soil yet. Your soil Hg measurements (0–30 cm depth-averaged) integrate over longer timescales, but the seasonal correlation (Fig. 6a) uses litter standing stock, not annualized litter Hg flux to soil.

Add a sentence in the Discussion: "Our litter Hg standing stocks represent a snapshot; the actual annual transfer of Hg from litter to mineral soil may be lower than our dry-season values suggest, because some dry-season litter Hg remains undecomposed and may be lost via re-emission or leaching before incorporation."

Our response: *Thank you for your comment. We agree that the measured litter Hg standing stocks represent a snapshot and do not necessarily reflect the whole amount of Hg transferred to soils. We will therefore clarify that the annual transfer of Hg from litter to soil may be lower than suggested by the dry-season litter Hg standing stock, because some Hg in litter may be lost through re-emission or leaching before incorporation into the soil.*

Final Recommendation

Major revision required. The litter data add valuable seasonal dynamics, but the manuscript currently:

Uses "inputs" incorrectly for standing stock measurements (Comment #9).

Overrelies on a simple correlation (Fig. 6a) without controlling for decomposition, organic matter, or post-senescence losses (Comment #7).

Contains inconsistent sample size reporting (Comment #10).

Makes overconfident comparisons with global means (Comment #11).

Addressing these issues (particularly #7 and #9) is essential before acceptance. The other comments (#10–14) are minor but would improve the clarity and accuracy of the manuscript.