nature portfolio

Peer Review File

Permafrost tipping point triggered by warming-driven loss of old carbon

Corresponding Author: Professor Jinzhi Ding

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript titled as “Permafrost tipping point triggered by warming-driven loss of old carbon” reveals a tipping point of ecosystems transition from weak to strong carbon sources driven by critical warming-induced decoupling between carbon fixation and respiration across temperature gradients based on a five-year, multi-level warming experiment on the Tibetan Plateau. The fate of vast carbon stocks stored in permafrost regions across the Tibetan Plateau under global warming represents a critical yet poorly constrained scientific issue. Overall, this manuscript is well-written and had a comprehensive data analysis. I have some specific comments as follows:
Why is soil temperature universally used for temperature sensitivity calculations? Could Ra (ecosystem respiration) and GPP be more sensitive to air temperature? Is the linear mixed-effects model appropriate for quantifying temperature sensitivity given nonlinear responses? The application of linear mixed-effects models to characterize temperature sensitivity appears problematic. Both GPP and Reco demonstrably exhibit nonlinear responses to warming. Linear approximations inherently fail to capture these critical dynamics.
The repeated assertion that 2–4°C warming triggers 'irreversible' carbon loss requires critical reevaluation. In fact, this five-year warming experiment cannot capture long-term soil carbon dynamics. Crucially, potential adaptive responses—such as vegetation shifts toward heat-tolerant communities and microbial functional resilience—remain unaccounted for in this temporal framework.
While the 'depletion of labile carbon substrates' is invoked to explain reduced Reco sensitivity under +4°C warming, this inference relies solely on flux dynamics and lacks direct chemical characterization of substrate pools (e.g., DOC concentrations). Have the authors ever considered the following mechanisms: Was dissolved organic carbon (DOC) pool variability quantified to assess its contribution to deep-layer CO_2 via subsurface transport? How were confounding effects of microbial community restructuring (e.g., proliferation of oligotrophic taxa) systematically excluded? Could physical protection mechanisms (e.g., mineral association in permafrost carbon pools) delay substrate depletion beyond the experimental timeframe? Another issue is that this study mainly considered CO2 variations without taking into account the changes in CH4.
L41-42: This sentence is a little bit confusing. Is respiration from by old carbon release from thawing permafrost also included in Reco?
L415: More details about the observation data need to be added, such as the missing data percentage, which determines whether your interpolation is effective.
L417：Why were soil temperature and moisture specifically at 10 cm and 20 cm depths selected?
L418-420: Is the same Random Forest model (using 10-20 cm soil data) applied to interpolate nocturnal NEE?
L421: The nighttime-based partitioning approach may introduce artificial interdependence between ecosystem respiration (Reco) and gross primary productivity (GPP) estimates.
Fig.2g- It is unclear what approach was used here to determine the relative contributions of Reco and GPP variations to alterations in NEE.

Reviewer #2

(Remarks to the Author)
This manuscript presents a multi-year, multi-level warming experiment in the Tibetan Plateau permafrost zone, integrating >40,000 hourly CO_2 flux measurements with depth-resolved $\delta^{13}C$-CO_2 profiling to identify critical temperature thresholds for

permafrost carbon balance and the mobilization of deep, old carbon. The study demonstrates a tipping point around 2 °C warming, beyond which GPP declines while Reco remains high, driven largely by deep carbon sources. The work is timely, methodologically innovative, and of high relevance to understanding permafrost climate feedbacks in Earth system models. However, I have several concerns that the authors should address before the manuscript can be accepted.

Major concerns:
1. The authors suggest that the attenuated respiratory sensitivity under prolonged +4 °C warming likely reflects substrate depletion in labile carbon pools. I question whether substantial substrate depletion can occur in such a short time frame. Direct evidence should be provided from soil carbon content measurements, not inferred solely from respiratory flux patterns.
2. While $\delta^{13}C$ signatures and radiocarbon data provide suggestive evidence, attributing elevated deep CO_2 concentrations solely to decomposition of old carbon may be problematic. Gas accumulation due to diffusion constraints or mixing effects could also contribute. The O_2 data partially address this, but the sampling frequency, seasonal coverage, and sensitivity may not be sufficient to fully rule out alternative explanations.
3. To ensure that the Bayesian stable isotope mixing model you use converges, you need to show the results of your final constraints and prove that you have enough information to constrain these components.
4. The manuscript reports that the decoupling between Reco and GPP drove a qualitative transition to a strong carbon source, implying a tipping point around 2 °C. However, the temperature gradient in the experiment is uneven, with relatively large gaps between treatments. While I understand that high-density continuous gradients are logistically challenging, the definition of thresholds in such cases should be more cautious. As these thresholds are used for future projections, the associated uncertainty could be very large and should be explicitly discussed.

Minor issues
1. Line 86, delete ~
2. Clarify why the results in Fig. 3e and 3f appear contradictory for the 1 °C and 2 °C treatments.
3. The unit of GPP is incorrect; as a rate, it should include a time component.
4. In Fig. 5, under Phase III, clarify why the absolute value of ΔGPP is greater than the absolute value of ΔER.
5. Line 134(also other places), change “oC” to “°C” in the unit g CO_2 m^{-2} d^{-1} $°C^{-1}$.
6. Maintain consistent use of significance symbols across figures (lowercase letters vs. asterisks) and explain the symbol scheme in each figure legend.
7. Double check the unit of GPP and ER in the manuscript are g CO2 not g C, usually GPP in QTP is about 200 to 900 g C m^{-2} yr^{-1}, if it is g CO2, it should be much larger.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have done a great job addressing all the reviewer’s comments and were very considerate, detailed and thoughtful in their responses. It seems that the manuscript has clearly benefited from the reviewer’s comments as it is much easier to understand, more clearly written and focused and more direct evidence. I just have several minor comments:
L60-To be consistent, cm a-1?
L457-The target warming magnitude is +1 °C, +2 °C, and +4 °C. So what is the actual temperature increase, and should be indicated and clarified.
It is recommended to strengthen the link between the Discussion and the two scientific hypotheses presented in the Introduction. For example, are the hypotheses supported? Why or why not?
Supplementary Fig. 27- all R^2 is equal to 1? These indicators require more explanation for understanding. Same for Supplementary Fig. 28

Reviewer #2

(Remarks to the Author)

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have addressed all my concerns. I have no more comments. Thanks for your great revision.

RESPONSE TO REVIEWERS' COMMENTS

We sincerely appreciate the two reviewers' great efforts in reviewing our manuscript and, especially, the constructive comments and insightful suggestions for further improving the manuscript's quality. Following is our point-by-point response to the reviewers, and the manuscript has been updated accordingly (with changes marked in *red*). For clarity and easier navigation, referee comments are presented in blue, while our responses are in black.

Reviewer #1

This manuscript titled as "Permafrost tipping point triggered by warming-driven loss of old carbon" reveals a tipping point of ecosystems transition from weak to strong carbon sources driven by critical warming-induced decoupling between carbon fixation and respiration across temperature gradients based on a five-year, multi-level warming experiment on the Tibetan Plateau. The fate of vast carbon stocks stored in permafrost regions across the Tibetan Plateau under global warming represents a critical yet poorly constrained scientific issue. Overall, this manuscript is well-written and had a comprehensive data analysis. I have some specific comments as follows:

Response: We appreciate your positive evaluation, kind encouragements and constructive comments on our work. We have carefully addressed all the issues you raised and revised the manuscript accordingly. Below, we provide point-by-point responses to all comments, and we hope our revision is to your satisfaction.

[Comment 1] Why is soil temperature universally used for temperature sensitivity calculations? Could Ra (ecosystem respiration) and GPP be more sensitive to air temperature?

Response: Thanks for this important question. Our use of soil temperature for calculating temperature sensitivity is based on its closer linkage to the key processes controlling both gross primary productivity (*GPP*) and ecosystem respiration (R_{eco}). While air temperature can fluctuate rapidly with weather and radiative conditions (**Fig. R1**), and often decouples from soil conditions, soil temperature provides a more stable and integrative measure of the thermal environment experienced by plants and soil microorganisms **under different warming treatments**.

For *GPP*, although canopy-level photosynthesis is driven by light and air temperature, the belowground environment exerts strong control on plant function. Soil temperature and moisture

determine root activity, nutrient mineralization, and water uptake, which ultimately constrain photosynthetic carbon assimilation. Thus, soil temperature reflects the integrated effects of both above- and belowground conditions on *GPP* more effectively than air temperature alone.

For R_{eco}, the belowground component (root and microbial respiration) constitutes the vast majority (~80%; **Fig. R2**) of total R_{eco} in our alpine meadow ecosystem and is directly governed by soil temperature. While aboveground plant respiration responds to air temperature, the overall metabolic activity of plants, including roots, is primarily constrained by the soil thermal regime. Therefore, soil temperature represents a more proximal and physiologically relevant driver of ecosystem-level carbon fluxes, offering a consistent basis for evaluating their temperature sensitivity under different warming treatments.

Taken together, we believe soil temperature provides a more appropriate and stable metric in our alpine meadow for comparing the temperature sensitivity of key carbon cycling processes under different warming treatments. Thanks for your understanding.

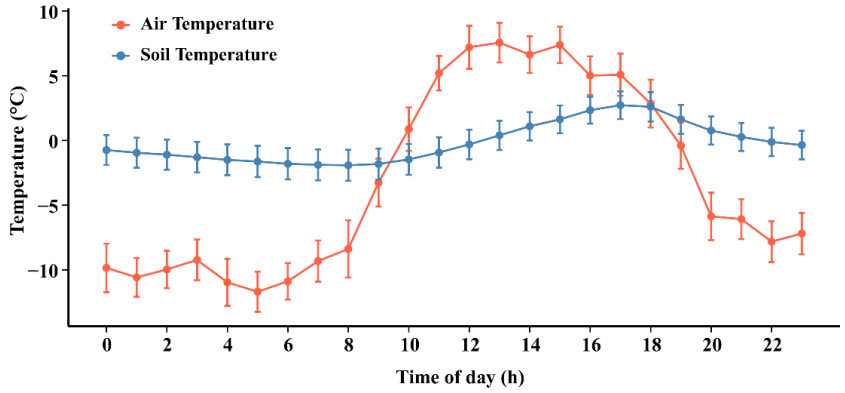

Fig. R1. The diurnal variations of air temperature and soil temperature (mean ± standard error).

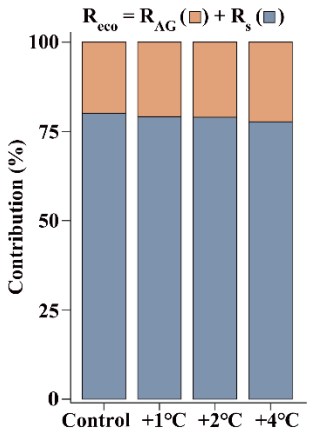

Fig. R2 (Fig. 4c). Contribution of aboveground plant respiration (R_{AG}) and soil respiration (R_S) to ecosystem respiration (R_{eco}).

[Comment 2] Is the linear mixed-effects model appropriate for quantifying temperature sensitivity given nonlinear responses? The application of linear mixed-effects models to characterize temperature sensitivity appears problematic. Both GPP and Reco demonstrably exhibit nonlinear responses to warming. Linear approximations inherently fail to capture these critical dynamics.

Response: Many thanks for this insightful comment. Ecosystem carbon processes indeed exhibit nonlinear responses to warming. In response, we have conducted additional analyses using nonlinear models (exponential functions or generalized additive mixed models (GAMMs)) to better characterize the temperature sensitivity of carbon fluxes. Importantly, the nonlinear analyses confirmed our original findings: the key patterns, thresholds, and conclusions remain unchanged (**Fig. R3**). This demonstrates our results are robust regardless of the modeling approach used. We have replaced the previous linear mixed-effects model results with those derived from the more appropriate nonlinear analysis in the updated manuscript.

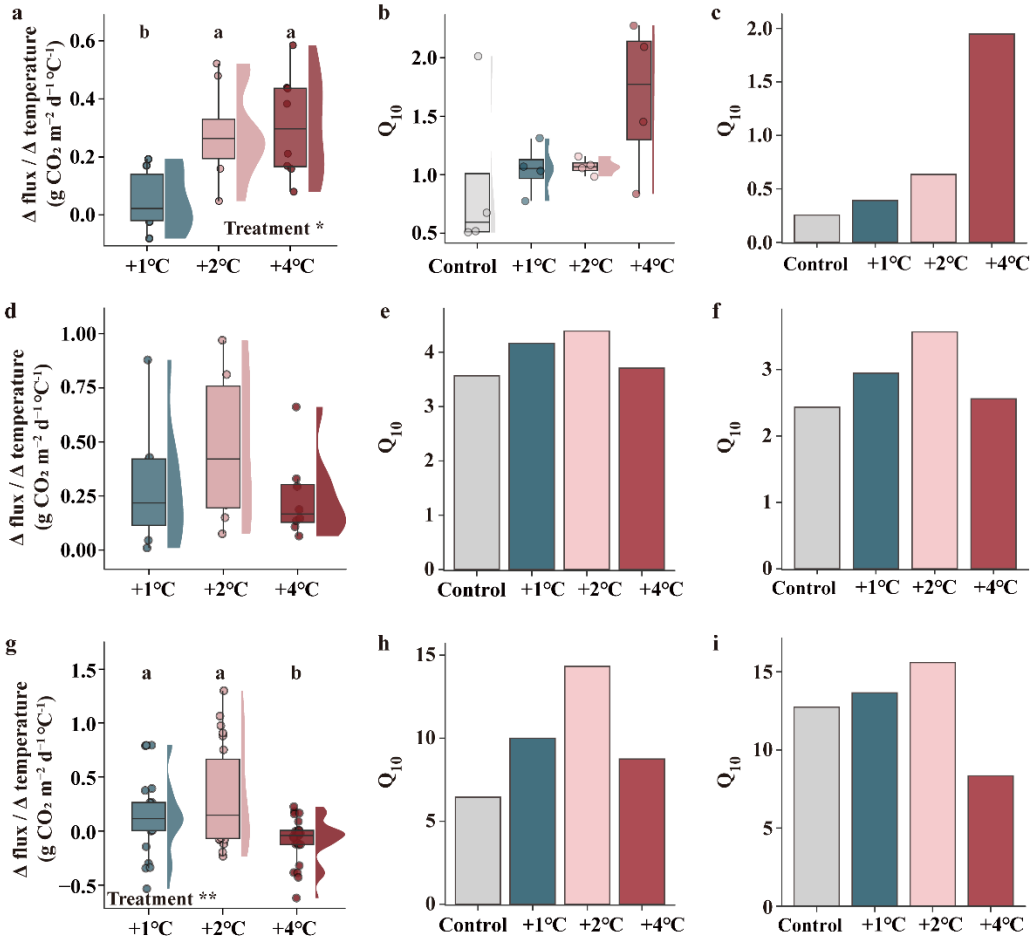

Fig. R3. Temperature sensitivity of net ecosystem CO_2 exchange (*NEE*, **a-c**), ecosystem respiration (R_{eco}, **d-f**), and gross primary productivity (*GPP*, **g-i**). **a, d, g,** Temperature sensitivity represented by change in flux per unit temperature change following multi-level warming. All

lowercase letters indicate significant differences between warming treatments. Significant level *$P < 0.05$ and **$P < 0.01$. **b, e, h,** Temperature sensitivity (Q_{10}) was quantified using non-linear models (exponential functions). Boxes denote interquartile range (25th-75th percentiles), central bars mark medians, whiskers extend to 1.5-fold the interquartile range (in the presence of outliers) or the minimum and maximum values (in the absence of outliers). The violin plot to the right of the box plot shows the density and distribution of the data. **c, f, i,** Temperature sensitivity (Q_{10}) was quantified by isolating the net warming effect with non-linear models (exponential functions or generalized additive mixed models (GAMMs)) that incorporated random effects and controlled for soil moisture covariation. GAMMs were used for R_{eco}, while exponential functions were used for *NEE* and *GPP*.

We have updated the calculation method of temperature sensitivity in the Methods section accordingly as follows: "*The temperature sensitivity of CO_2 fluxes was quantified using three complementary approaches. First, we calculated the direct sensitivity metric, defined as the change in flux (Δflux) per 1 °C increase relative to the control, to represent the net warming effect under different treatments. Second, we assessed the temporal temperature sensitivity (Q_{10}) using exponential functions by analyzing the time series relationships between fluxes and temperature across warming levels, as follows (Eq. 2, Eq. 3[62]):*

$$R = ae^{bt} \quad (2)$$

$$Q_{10} = e^{10b} \quad (3)$$

where R represents CO_2 flux, a denotes the flux at 0 °C, b denotes the model fitting parameter, t denotes the soil temperature at the 10 cm layer, and Q_{10} indicates the temperature sensitivity, that is, the multiplicative increase in the CO_2 flux for every 10 °C increase in soil temperature. Third, because humidity fluctuations can influence temperature responses, we conducted partial correlation analyses using nonlinear models (exponential functions or generalized additive mixed models (GAMMs)) controlling for humidity to isolate and quantify the temperature-driven component of flux variation. Using the fitted GAMMs, we derived a dynamic Q_{10} by calculating the flux ratio (Flux(T+10) / Flux(T)) for each 1 °C step within the observed temperature range. We then presented the average Q_{10} for each treatment over this range." (lines 632−646) We also made corresponding revisions in the Results and Discussion sections (lines 92−95, lines 131−138, lines 231−233, and lines 301−302). We believe this revision strengthens our study by employing more biologically realistic temperature response functions.

[Comment 3] The repeated assertion that 2–4°C warming triggers 'irreversible' carbon loss requires critical revaluation. In fact, this five-year warming experiment cannot capture long-term

soil carbon dynamics. Crucially, potential adaptive responses—such as vegetation shifts toward heat-tolerant communities and microbial functional resilience—remain unaccounted for in this temporal framework.

Response: We agree the reviewer that the term "irreversible" may be too strong for a five-year experiment and that our study cannot capture long-term soil carbon dynamics. Such a timeframe is also insufficient to capture the long-term adaptive responses of vegetation and microorganisms. To address this concern, we have made the following revisions:

(i)We softened the wording of "irreversible" carbon loss to more accurate and cautious expressions, such as "*persistent*" or "*potentially irreversible*" carbon loss, throughout the manuscript (line 53 and line 311).

(ii) We have explicitly framed our findings within the timeframe of our observation period, and revised the relevant statement as follows:

"*During the five-year experimental period, this study identifies a climate tipping point in an alpine ecosystem of the Tibetan permafrost region (within 2−4 °C warming), and reveals a critical warming-induced decoupling between carbon fixation and loss, driving the ecosystem transition from weak to strong carbon source across temperature gradients (Fig. 5).*" (lines 417−420)

(iii) We have added a new paragraph in the Discussion to discuss the potential for long-term ecosystem adaptations that could alter the trajectory observed. This includes vegetation shifts toward heat- and drought-tolerant communities and microbial functional adaptation, which represent important avenues for future research. The detailed revisions are as follows:

"*While our experiment documented a clear and persistent shift in the carbon cycling, its five-year duration necessarily limits inference about long-term trajectories. Potential adaptive responses, which are particularly relevant in the context of the vulnerable Tibetan Plateau alpine ecosystem, remain unaccounted for. For instance, plant communities may gradually shift from moisture-sensitive sedges (e.g., Kobresia tibetica and Kobresia pygmaea) towards more heat- and drought-tolerant grasses and forbs. Such transitions could partially stabilize GPP during dry periods, but also alter belowground carbon cycling through changes in root architecture and litter quality, potentially not fully compensating for the observed productivity collapse. Similarly, microbial communities may exhibit functional resilience through the enrichment of thermotolerant taxa or a shift towards oligotrophic groups (e.g., Actinobacteria) specialized in decomposing recalcitrant old carbon. This functional succession could reduce the temperature sensitivity of decomposition but concurrently lead to the long-term, persistent mineralization of the previously protected carbon pool. These potential acclimations represent critical*

uncertainties in projecting the permanence of the carbon loss we observed. Therefore, future work integrating long-term monitoring, vegetation dynamics, and microbial genomic analyses is crucial to test these hypotheses and determine the permanence of permafrost carbon loss we document." (lines 399−414)

We believe these revisions have significantly tempered our claims, improved the accuracy of our language, and positioned our study as documenting a critical short-term trajectory while acknowledging the complexities of long-term permafrost carbon dynamics.

[Comment 4] While the 'depletion of labile carbon substrates' is invoked to explain reduced Reco sensitivity under +4°C warming, this inference relies solely on flux dynamics and lacks direct chemical characterization of substrate pools (e.g., DOC concentrations). Have the authors ever considered the following mechanisms: Was dissolved organic carbon (DOC) pool variability quantified to assess its contribution to deep-layer CO_2 via subsurface transport? How were confounding effects of microbial community restructuring (e.g., proliferation of oligotrophic taxa) systematically excluded? Could physical protection mechanisms (e.g., mineral association in permafrost carbon pools) delay substrate depletion beyond the experimental timeframe?

Response: We are deeply grateful to the reviewer for these insightful comments. We fully agree that 'substrate depletion' cannot be inferred from flux dynamics alone and requires direct chemical and biological validation. **In response, we made substantial additional efforts, including new soil sampling, measurements, and analyses of soil carbon fractions, microbial metagenomes, and phospholipid-derived fatty acids (PLFAs).** Below, we address the specific points raised.

(1) **Direct chemical evidence for substrate depletion:** Warming by +4 °C significantly reduced particulate organic carbon (POC) of surface soils, a major labile carbon pool, providing direct chemical evidence of the depletion of labile substrates. Concurrently, mineral-associated organic carbon (MAOC) also declined significantly, indicating that warming mobilized carbon beyond the labile fraction and into mineral-stabilized pools, mechanistically explains the observed attenuation of temperature sensitivity in ecosystem respiration (**Fig. R4**).

(2) **Regarding the reviewer's question on potential dissolved organic carbon (DOC) transport**, we quantified DOC concentrations in both surface (0−10 cm) and deep (70−100 cm) soils to assess whether DOC redistribution could contribute to deep-layer CO_2 production. If substantial DOC leaching or downward transport had occurred, an enrichment of DOC in deep soils would have been expected. However, our measurements show that surface DOC exhibited

an increasing trend under +4 °C warming, consistent with enhanced plant-derived carbon inputs associated with increased *GPP*. In contrast, deep-soil DOC did not increase and instead showed a declining trend across warming treatments (**Fig. R4**), indicating no detectable downward transfer of dissolved carbon.

Therefore, the DOC data do not support the hypothesis that deep-layer CO_2 originates from DOC transport. Rather, the warming-induced CO_2 patterns in deep soils are best explained by direct microbial temperature responses acting on the deep carbon pool, with the decline in DOC providing evidence for enhanced deep-soil carbon decomposition. Although deep-layer POC and MAOC did not show significant changes over the five-year period, likely due to intrinsic variability, analytical noise, and the relatively short experimental duration, the deep-layer CO_2 production will inevitably lead to a decline in the deep-soil SOC pool over longer timescales.

(3) **Microbial community restructuring is mechanistically coupled to substrate depletion, rather than confounding our interpretation.** Metagenomic analyses revealed a functional and taxonomic shift consistent with carbon limitation under +4 °C warming treatment (**Fig. R5**). Specifically, warming favored oligotrophic taxa adapted to low-carbon environments, such as *Candidatus_Dormibacterota*, known for dormancy strategies, and members of *Verrucomicrobiota* capable of degrading chemically complex substrates[1], while other oligotrophic groups (e.g., *Acidobacteriota*) declined. This pattern reflects a selective enrichment of taxa pre-adapted to substrate scarcity, aligning with the depletion of labile carbon pools observed above. Such microbial community restructuring represents a biological response to substrate depletion, contributing to the observed reduction in temperature sensitivity.

PLFAs data further support this interpretation. The fungi-to-bacteria ratio (F : B) declined significantly under +4 °C warming (**Fig. R6**), indicating a shift toward bacterial dominance typical of carbon-limited conditions. Although changes in the Gram-positive to Gram-negative ratio (G^+ : G^-) and monounsaturated fatty acids (MUFA, indicative of G^- bacteria) were not statistically significant, their synergetic trends (an increasing G^+ : G^- and decreasing MUFA) point to a microbial community restructuring toward stress-tolerant, resource-conservative groups, suggesting reduced carbon availability rather than enhanced microbial turnover.

(4) **Regarding physical protection mechanisms**, for the surface layer, MAOC decreased substantially under the +4 °C warming treatment, indicating that the mineral-protected carbon pool was not sufficiently stable to buffer substrate depletion within the five-year timescale of our experiment. For the deep layer, we did not observe detectable changes in MAOC. We agree that mineral association can partially slow the depletion of deep carbon, which may contribute to the absence of detectable MAOC shifts within five years. However, this is not the only

explanation. The warming signal at the surface (+4 °C) attenuated to <1 °C in deep soils, and deep mineral-associated carbon inherently turns over much more slowly than surface pools. Together, these factors make short-term changes in deep MAOC difficult to detect, even if decomposition processes are occurring.

We have made corresponding revisions in the Methods section (lines 487−526), and integrated this multi-faceted evidence into the revised Discussion as follows:

"*Instead, the attenuated respiratory sensitivity under prolonged +4 °C warming reflects substrate depletion in labile carbon pools and the subsequent functional succession of the microbial community. First, direct quantification of soil carbon fractions revealed a significant reduction in particulate organic carbon (POC; Supplementary Fig. 15), a major labile carbon pool, in surface soils, providing direct chemical evidence of the depletion of labile substrates. Surface DOC, which comprises <1% of total SOC, showed an increasing trend associated with warming-enhanced plant inputs, yet this small increase was insufficient to counterbalance the substantial POC loss. Furthermore, warming-induced decomposition was not confined to the labile fraction. MAOC, regarded as chemically recalcitrant pool, also decreased significantly in surface soils (Supplementary Fig. 15), indicating that accelerated microbial activity had begun to mobilize this protected carbon pool. These changes collectively resulted in a significant reduction in total SOC. Deep-soil DOC also exhibited a clear decline of 26%, indicating substantial losses of labile carbon at depth. Although deep-layer POC and MAOC did not show statistically significant changes, likely due to intrinsic variability, analytical noise, and the much weaker warming signal (<1 °C) transmitted to deeper layers, the persistently high CO_2 production from these horizons nonetheless points to enhanced decomposition of deep soil organic matter.*

Second, the temporal pattern of respiration is fully consistent with this mechanism: initial linear R_{eco} increases in 2021 (Supplementary Fig. 16) gave way to nonlinear declines (2022−2024) as labile carbon pools became depleted, especially under +4 °C warming. Critically, this substrate-depletion scenario is further corroborated by a fundamental restructuring of the soil microbial community. Metagenomic and lipid biomarker analyses revealed a systematic microbial shift toward oligotrophic life-history strategies under +4 °C warming. This transition was marked by the enrichment of taxa adapted to resource scarcity (Supplementary Fig. 17), such as Candidatus_Dormibacterota (employing dormancy strategies) and Verrucomicrobiota (capable of complex carbon degradation)[42]. Concurrently, lipid profiles showed a significant decline in the fungal-to-bacterial ratio (F : B, a recognized biosignature of carbon limitation), accompanied by concerted shifts in other stress indicators (an increasing

Gram-positive to Gram-negative ratio (G+ : G-) and decreasing monounsaturated fatty acids (MUFA); Supplementary Fig. 18), collectively pointing to a microbial community restructuring toward stress-tolerant, resource-conservative groups, suggesting reduced carbon availability rather than enhanced microbial turnover. Collectively, these findings establish a coherent biogeochemical framework: intense warming progressively depletes labile substrates, which in turn drives a microbial community shift toward resource-conservative strategies, ultimately attenuating the ecosystem-scale respiratory response." (lines 254−287)

"*We initially hypothesized that dissolved organic carbon (DOC) transport from surface layers could fuel this deep-layer CO_2 production. However, direct quantification of DOC concentrations refutes this mechanism: while surface DOC exhibited an increasing trend under +4 °C warming (consistent with enhanced plant-derived inputs), deep-soil DOC concurrently declined (Supplementary Fig. 15), showing no evidence of substantial downward carbon transfer.*" (lines 337−341)

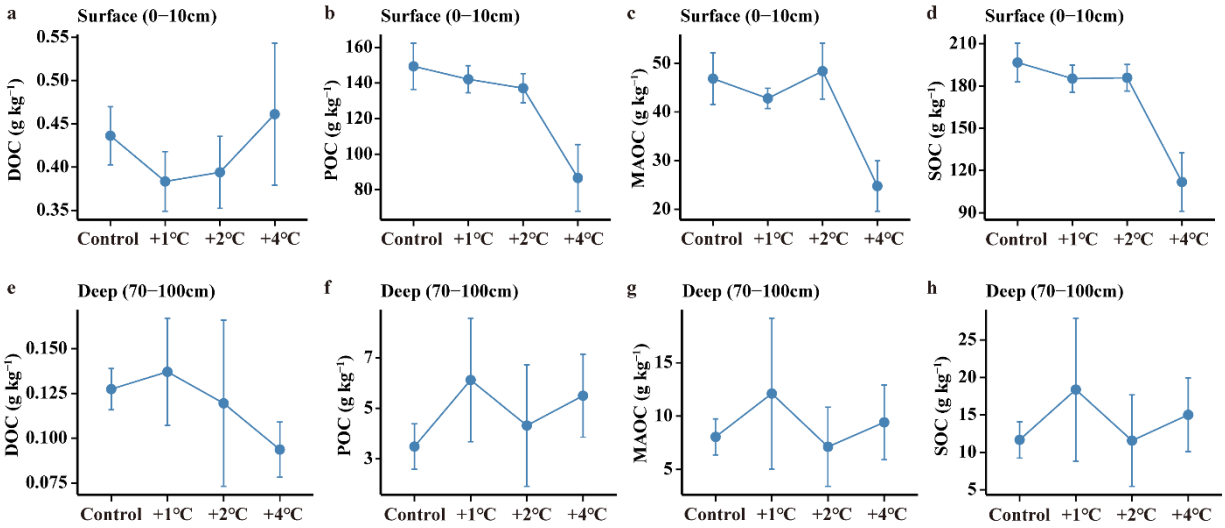

Fig. R4 (Supplementary Figure 15). Effects of warming gradients on dissolved organic carbon (DOC), particulate organic carbon (POC), mineral-associated organic carbon (MAOC), and soil organic carbon (SOC) in surface (0–10cm, **a-d**) and deep soil layers (70–100cm, **e-h**) (mean ± standard error).

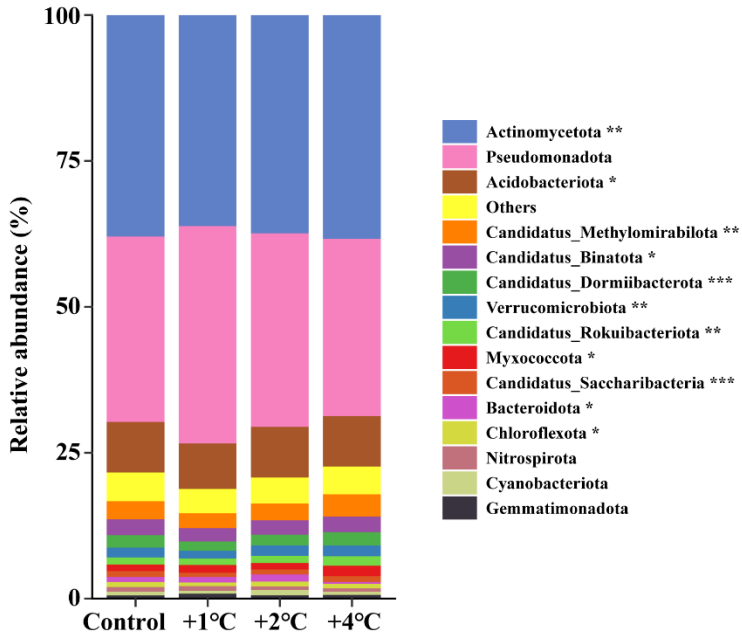

Fig. R5 (Supplementary Figure 17). Effects of warming gradients on microbial community. The relative abundances at phylum level. Based on ANOVA, bacterial phyla with significant differences across warming treatments were marked with '*'. Significance levels were as follows: *P < 0.05, **P < 0.01, and ***P < 0.001.

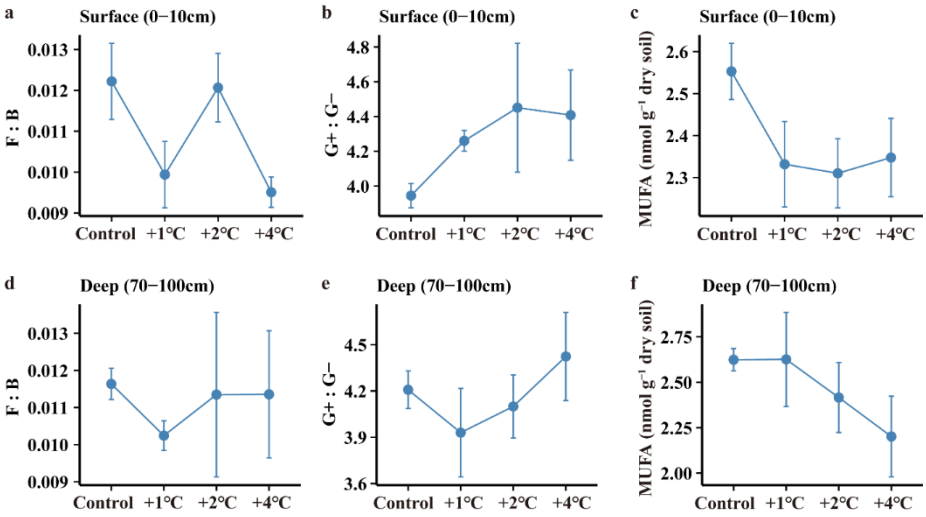

Fig. R6 (Supplementary Figure 18). Effects of warming gradients on the fungal-to-bacterial ratio (F : B), Gram-positive and Gram-negative ratio (G+ : G-), and monounsaturated fatty acids (MUFA) PLFAs in surface (0–10cm, **a-c**) and deep soil layers (70–100cm, **d-f**) (mean ± standard error).

[Comment 5] Another issue is that this study mainly considered CO2 variations without taking into account the changes in CH4.

Response: In fact, we measured methane (CH_4) fluxes concurrently throughout the entire experimental period. Our analysis reveals that the alpine meadow ecosystem functioned as a CH_4 sink, and gradient warming enhanced this annual CH_4 uptake, with a significant effect observed under high-warming treatment (**Fig. R7**). To evaluate the climate impact of CH_4 uptake in the context of the overall greenhouse gas balance, we converted CH_4 uptake to CO_2-equivalents using GWP-20[2]. The resulting offsets accounted for only 0.92−1.85% of the substantial CO_2 emissions (**Fig. R7**). This indicates that CH_4 uptake plays a minor role in the overall greenhouse gas balance, and the ecosystem's climate feedback is overwhelmingly dominated by CO_2 fluxes. Therefore, inclusion of CH_4 does not alter our central conclusion that sustained warming triggers a carbon-climate tipping point primarily through CO_2 dynamics. We plan to report the full CH_4 dataset in a separate manuscript that focuses specifically on non-CO_2 greenhouse gas responses to warming. Thanks for your understanding.

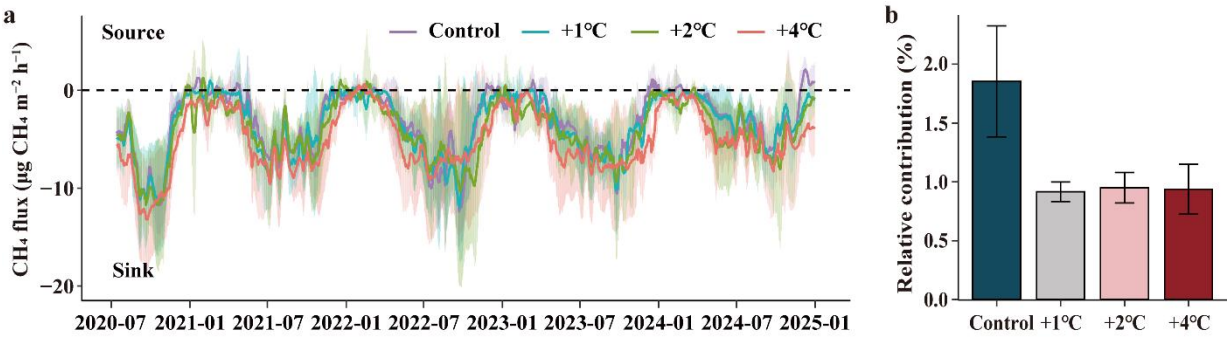

Fig. R7. Multi-gradient warming effects on ecosystem CH_4 uptake over 2020−2024 (**a**) and contribution of CH_4 absorption to CO_2 emissions (**b**). The shaded area different warming treatments represents the mean ± standard deviation. Error bars represent the mean ± standard error.

[Comment 6] L41-42: This sentence is a little bit confusing. Is respiration from by old carbon release from thawing permafrost also included in Reco?

Response: In fact, the decomposition of old carbon is part of R_{eco}. To avoid any potential misunderstanding, we have rewritten the sentence to clarify that the third dynamic is a shift within R_{eco} itself—from the decomposition of recent photosynthate to the mineralization of old carbon from thawing permafrost:

"*The determination of permafrost carbon-climate thresholds depends on three temperature-related dynamics, including carbon gain driven by warming-enhanced gross primary productivity (GPP), carbon loss via ecosystem respiration (R_{eco}) acceleration, and a pivotal shift within R_{eco} from decomposing recent photosynthate to mineralizing old carbon.*" (lines 38−41)

[Comment 7] L415: More details about the observation data need to be added, such as the missing data percentage, which determines whether your interpolation is effective.

Response: Thanks for this nice suggestion. In the revised manuscript, we reported that approximately 30% of the hourly *NEE* data were missing, primarily due to instrument malfunction and power outages. These gaps were filled using a Random Forest regression model, chosen for its strong ability to represent non-linear CO_2-environment relationships. Model performance was validated by predicting masked observations, and the results confirmed that the gap-filling was reliable. Details and validation plots are provided in the revised Methods (lines 538−552) and Supplementary Fig. 24.

[Comment 8] L417: Why were soil temperature and moisture specifically at 10 cm and 20 cm depths selected?

Response: Our selection of soil temperature and moisture at 10 cm and 20 cm depths was based on ecological relevance and model performance. These depths represent the primary root zone and the layer of highest microbial activity in alpine meadows, and thus best capture the belowground conditions governing both *GPP* and R_{eco}. Moreover, soil temperature and moisture integrate environmental constraints more effectively than air temperature and humidity, particularly because belowground respiration contributes ~80% of total R_{eco} in this ecosystem (see response to Comment 1 for details).

We also tested alternative predictor combinations and found that using 10 cm and 20 cm data provided the optimal balance between predictive accuracy and model parsimony; including additional depths did not substantially improve model performance. Relevant clarifications have been added to the revised manuscript as follows:

"*Soil temperature and moisture at 10 cm and 20 cm depths were selected as RF model predictors because they directly reflect the root zone and the layer of highest microbial activity in alpine meadows. This combination provided the simplest model while outperforming models based on air temperature or other soil depths.*" (lines 544−547)

[Comment 9] L418-420: Is the same Random Forest model (using 10-20 cm soil data) applied to interpolate nocturnal NEE?

Response: Yes. The same Random Forest model, trained on the relationship between nocturnal *NEE* (which equals R_{eco}) and soil temperature and moisture at 10 cm and 20 cm depths, was applied to estimate diurnal R_{eco}. This approach is justified because R_{eco} is dominated by belowground microbial and root respiration, which is directly regulated by the soil environment

(see responses to Comment 1 and Comment 8 for details). The strong predictive performance of the model further confirms that these soil variables reliably capture R_{eco} dynamics across the full diel cycle. We have updated methods description as follows:

"*The same RF model, trained on the relationship between nocturnal NEE (which equals R_{eco}) and soil temperature and moisture at 10 cm and 20 cm depths, was applied to estimate diurnal R_{eco}, resulting in a final model with R^2 values ranging from 0.79 to 0.92 (Supplementary Fig. 25).*" (lines 552−554)

[Comment 10] L421: The nighttime-based partitioning approach may introduce artificial interdependence between ecosystem respiration (Reco) and gross primary productivity (GPP) estimates.

Response: To ensure that the nighttime-based partitioning approach did not bias our results, we conducted an independent girdling experiment that provides direct measurements of R_{eco} and is not subject to this interdependence[3]. We found that these independent measurements showed strong agreement with the R_{eco} values predicted by our Random Forest model (r = 0.88, $P < 0.001$; **Fig. R8**). This validation demonstrates that our R_{eco} estimates are robust and that any propagation of model error into *GPP* is minimal.

As girdling cannot provide continuous time series required for threshold detection, we used nighttime-based partitioning for generating high-frequency *GPP* and R_{eco}, while relying on independent girdling measurements to confirm the reliability of the partitioning approach.

In response to the reviewer's comment, we have explicitly acknowledged the potential interdependence introduced by the nighttime-based partitioning method in the Methods section as follows:

"*We note that the nighttime-based partitioning method, which can introduce artificial interdependence between GPP and R_{eco} estimates. To independently validate our R_{eco} values, we conducted girdling experiments that directly measure R_{eco} and are free from this interdependence*[30]*. The strong agreement between girdling-derived and model-derived R_{eco} provides confidence in the reliability of our partitioned fluxes (Supplementary Fig. 26).*" (lines 569−573)

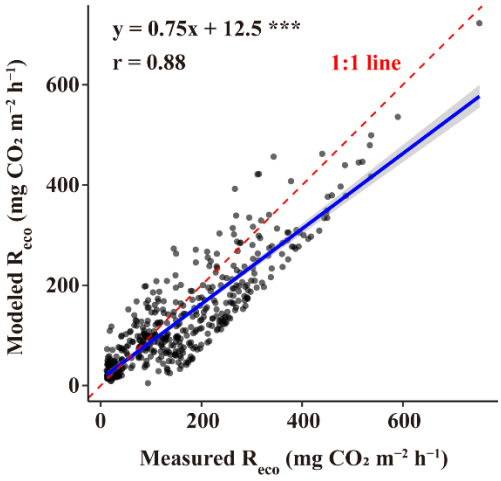

Fig. R8 (Supplementary Figure 26). Comparison of measured and modeled R_{eco}. The blue lines indicate regression fits, the gray shadow indicates the 95% confidence interval of the fitting line, and the fitting equation, r, and significance (***P < 0.001) are annotated. The dashed lines represent 1:1 line.

[Comment 11] Fig.2g- It is unclear what approach was used here to determine the relative contributions of Reco and GPP variations to alterations in NEE.

Response: Our analysis in Fig. 2g evaluates the relative impact strength of R_{eco} and GPP changes on NEE alterations using following metric:

$$\text{Impact strength} = |\Delta\text{Component}| / (|\Delta R_{eco}| + |\Delta GPP|)$$

This metric quantifies which process had the larger absolute effect on driving NEE changes, focusing on the magnitude of influence rather than directional effects. We have updated the figure legend explicitly clarified this point: "*Values represent the proportional influence of each component's absolute change magnitude to relative to the combined absolute changes of R_{eco} and GPP on the overall NEE response.*" (lines 123−124)

We have also included a detailed description of the calculation procedure in the Methods section as follows:

"*To assess the relative influence of R_{eco} and GPP variations on NEE changes, we calculated the proportional impact strength of each component:*

$$\textit{Impact strength of } R_{eco} = |\Delta R_{eco}| / (|\Delta R_{eco}| + |\Delta GPP|)$$

$$\textit{Impact strength of GPP} = |\Delta GPP| / (|\Delta R_{eco}| + |\Delta GPP|)$$

This approach quantifies which process had the larger absolute effect on altering the net carbon balance, regardless of the directional influence." (lines 562−567)

Reviewer #2

General Comment: This manuscript presents a multi-year, multi-level warming experiment in the Tibetan Plateau permafrost zone, integrating >40,000 hourly CO_2 flux measurements with depth-resolved $\delta^{13}C$-CO_2 profiling to identify critical temperature thresholds for permafrost carbon balance and the mobilization of deep, old carbon. The study demonstrates a tipping point around 2 °C warming, beyond which GPP declines while Reco remains high, driven largely by deep carbon sources. The work is timely, methodologically innovative, and of high relevance to understanding permafrost climate feedbacks in Earth system models. However, I have several concerns that the authors should address before the manuscript can be accepted.

Response: Thank you very much for your supportive evaluation and constructive comments that have substantially improved our manuscript. We have carefully addressed all concerns raised and revised the paper accordingly. We believe these revisions have significantly strengthened the clarity and robustness of the work.

Major concerns:

[Comment 1] The authors suggest that the attenuated respiratory sensitivity under prolonged +4 °C warming likely reflects substrate depletion in labile carbon pools. I question whether substantial substrate depletion can occur in such a short time frame. Direct evidence should be provided from soil carbon content measurements, not inferred solely from respiratory flux patterns.

Response: Thank you for raising this important mechanistic question. We agree that substrate depletion should be supported by direct soil carbon measurements rather than inferred solely from respiratory patterns. **In response, we made substantial additional efforts, including new soil sampling, measurements, and analyses of soil carbon fractions.** We quantified dissolved organic carbon (DOC) and particulate organic carbon (POC), which together represent the dominant labile carbon pools in this ecosystem. We also measured mineral-associated organic carbon (MAOC) to provide a more comprehensive assessment of changes across major soil carbon fractions (**Fig. R9**).

Our results provide direct evidence for substrate depletion under +4 °C warming. Surface POC, which constitutes the major fraction of the labile substrate pool, declined significantly within five years experiment. Surface DOC, which comprises <1% of total SOC, showed an increasing trend associated with warming-enhanced plant inputs, yet this small increase was insufficient to counterbalance the substantial POC loss.

Furthermore, warming-induced decomposition was not confined to the labile fraction. MAOC, regarded as chemically recalcitrant pool, also decreased significantly in surface soils, indicating that accelerated microbial activity had begun to mobilize this protected carbon pool. These changes collectively resulted in a significant reduction in total SOC. Deep-soil DOC also exhibited a clear decline of 26%, indicating substantial losses of labile carbon at depth. Although deep-layer POC and MAOC did not show statistically significant changes, likely due to intrinsic variability, analytical noise, and the much weaker warming signal (<1 °C) transmitted to deeper layers, the persistently high CO_2 production from these horizons nonetheless points to enhanced decomposition of deep soil organic matter. Together, these new soil carbon fraction data provide the direct chemical evidence requested by the reviewer and strongly support our interpretation that substrate depletion is a key mechanism underlying the attenuated respiratory sensitivity under +4 °C warming.

We have integrated this multi-faceted evidence into the revised Discussion as follows:.

"*Instead, the attenuated respiratory sensitivity under prolonged +4 °C warming reflects substrate depletion in labile carbon pools and the subsequent functional succession of the microbial community. First, direct quantification of soil carbon fractions revealed a significant reduction in particulate organic carbon (POC; Supplementary Fig. 15), a major labile carbon pool, in surface soils, providing direct chemical evidence of the depletion of labile substrates. Surface DOC, which comprises <1% of total SOC, showed an increasing trend associated with warming-enhanced plant inputs, yet this small increase was insufficient to counterbalance the substantial POC loss. Furthermore, warming-induced decomposition was not confined to the labile fraction. MAOC, regarded as chemically recalcitrant pool, also decreased significantly in surface soils (Supplementary Fig. 15), indicating that accelerated microbial activity had begun to mobilize this protected carbon pool. These changes collectively resulted in a significant reduction in total SOC. Deep-soil DOC also exhibited a clear decline of 26%, indicating substantial losses of labile carbon at depth. Although deep-layer POC and MAOC did not show statistically significant changes, likely due to intrinsic variability, analytical noise, and the much weaker warming signal (<1 °C) transmitted to deeper layers, the persistently high CO_2 production from these horizons nonetheless points to enhanced decomposition of deep soil organic matter.*" (lines 254−269)

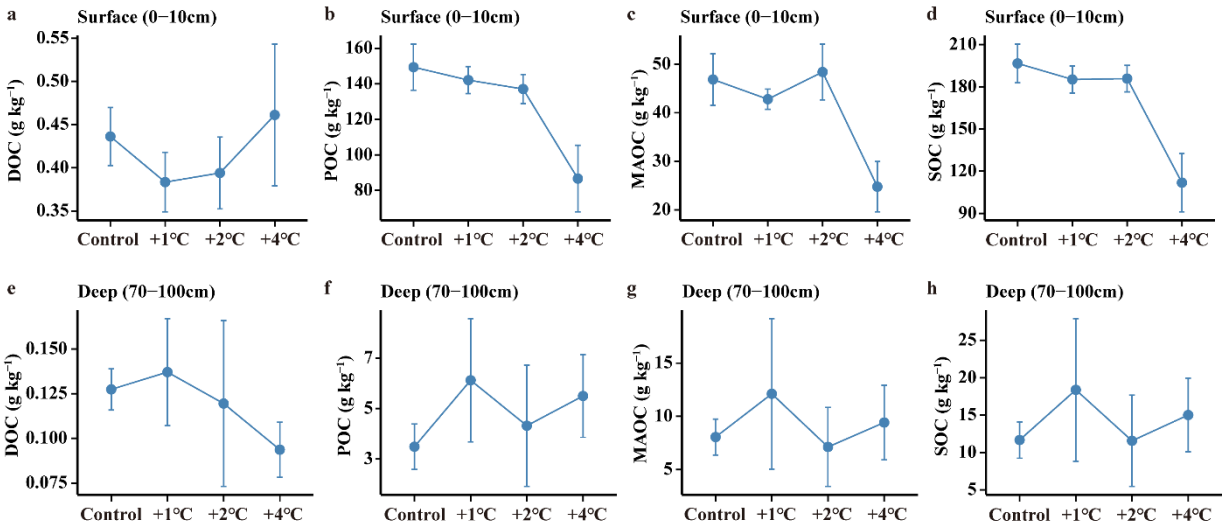

Fig. R9 (Supplementary Figure 15). Effects of warming gradients on dissolved organic carbon (DOC), particulate organic carbon (POC), mineral-associated organic carbon (MAOC), and soil organic carbon (SOC) in surface (0–10cm, **a-d**) and deep soil layers (70–100cm, **e-h**) (mean ± standard error).

[Comment 2] While $\delta^{13}C$ signatures and radiocarbon data provide suggestive evidence, attributing elevated deep CO_2 concentrations solely to decomposition of old carbon may be problematic. Gas accumulation due to diffusion constraints or mixing effects could also contribute. The O_2 data partially address this, but the sampling frequency, seasonal coverage, and sensitivity may not be sufficient to fully rule out alternative explanations.

Response: Very insightful comment. Isotopic evidence alone cannot conclusively attribute elevated deep CO_2 concentrations to old-carbon decomposition. The reviewer highlights two key alternatives: (i) CO_2 accumulation caused by diffusion constraints and (ii) gas mixing effects independent of biological mineralization.

To directly assess the diffusion-constraint hypothesis, we reanalyzed the soil O_2 profiles for different seasons based on data collected from March to November. Measurements were made using a high-precision optical O_2 sensor (TD400-SH-O_2; accuracy ±0.18 vol%; repeatability 0.2%), which is capable of detecting subtle vertical gradients in well-aerated soils. Across different seasons, no statistically significant differences in O_2 concentrations were observed between surface and deep layers (**Fig. R10**). This result indicates that the soil column remained consistently well ventilated. A diffusion-limited system would necessarily exhibit deep O_2 depletion coupled with CO_2 accumulation, a pattern not supported by our data.

The reviewer also raises the possibility of mixing effects. We note that mixing should follow the natural concentration gradient: gas should diffuse from deep layers (high CO_2) to shallow layers (low CO_2). Reverse-gradient mixing, where deep CO_2 remains high despite continuous diffusive loss upward, would require physical barriers or sustained downward advection, neither of which is supported by soil structure or hydrological conditions at our site. Therefore, while minor mixing undoubtedly occurs, it cannot sustain the observed persistent high CO_2 at depth, especially in the absence of O_2 depletion. This again points to continuous in situ production rather than physical redistribution.

Taken together, while diffusion constraints and gas mixing may contribute to some variability, they are unlikely to be the primary drivers of the deep-soil CO_2 patterns. The well-ventilated soil profile indicated by the O_2 measurements, together with the isotopic evidence, points to old-carbon decomposition rather than physical trapping or mixing as the dominant source of high-level deep CO_2. We have incorporated these points into Discussion.

We agree with the reviewer that our O_2 dataset does not cover the entire frozen period, when diffusion may be restricted. However, for diffusion constraints to serve as the dominant explanation, strong O_2 depletion would need to persist for much of the year, which is inconsistent with our nine-month dataset. We also explicitly discuss this limitation in the revised manuscript as follows:

"*Another important alternative explanation for elevated deep-soil CO_2 is physical gas accumulation caused by diffusion constraints or mixing effects. However, analysis of oxygen (O_2) profiles resolved from March to November revealed no significant vertical gradients in different seasons (Supplementary Fig. 20), indicating effective soil ventilation during the entire period (soil moisture <40%; Supplementary Fig. 14). A diffusion-limited system would display persistent O_2 depletion coinciding with CO_2 accumulation, which was not observed. Instead, the combination of elevated deep-soil CO_2 and only modest O_2 declines at depth is most consistent with continuous in situ biological production, in agreement with the $\delta^{13}C$ and radiocarbon evidence for old-carbon decomposition. We acknowledge that our O_2 dataset does not extend into the entire frozen period, during which diffusion limitations may be enhanced. Nevertheless, such limitations would need to persist throughout much of the year to explain the observed CO_2 patterns, which is not supported by our multi-month measurements. Taken together, these lines of evidence suggest that while diffusion constraints may play a secondary role, the decomposition of old carbon is the predominant source of deep-soil CO_2 during the period investigated.*" (lines 341−355)

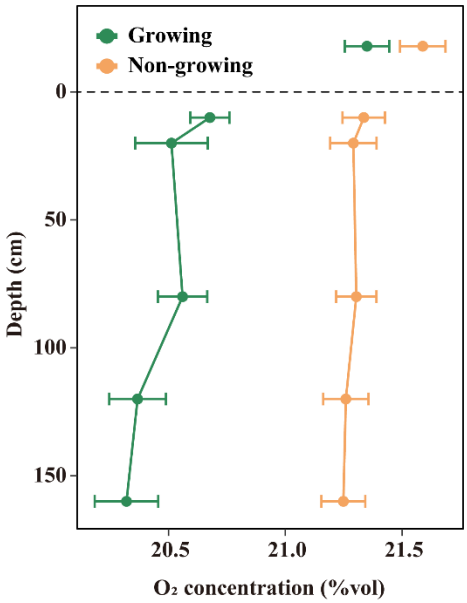

Fig. R10. Depth profiles of soil oxygen (O_2) concentration and atmospheric O_2 levels in the growing and non-growing seasons (mean ± standard error). Below the dashed line, points and lines represent soil profile O_2 concentration; above, atmospheric O_2 is shown.

[Comment 3] To ensure that the Bayesian stable isotope mixing model you use converges, you need to show the results of your final constraints and prove that you have enough information to constrain these components.

Response: Following this constructive suggestion, we have added comprehensive diagnostic outputs to support our interpretation:

(i) Model Convergence: Gelman-Rubin diagnostics confirm full convergence of all parameters, with $\hat{R}$ values < 1.01 (**Fig. R11**). Geweke diagnostics confirmed chain stationarity, with all z-scores within ±1.96 (**Fig. R12**). Trace plots demonstrate excellent mixing of the multiple chains with no trends (**Fig. R13**).

(ii) Additional Validation: Autocorrelation plots show rapid decay to zero, ensuring sample independence (**Fig. R14**).

(iii) Parameter Constraint: The posterior distributions demonstrate that the isotopic data provide meaningful information to constrain source contributions (**Fig. R15**). Although some posterior ranges are broad, reflecting the inherent difficulty of separating isotopically similar soil carbon pools, the clear peak locations and largely non-overlapping credible intervals indicate that the model successfully retrieves identifiable and interpretable source signals.

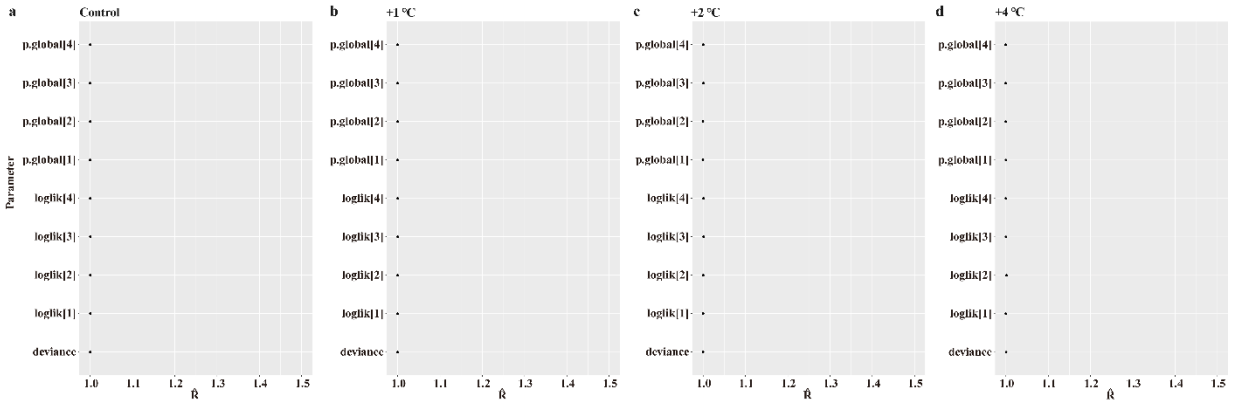

Fig. R11 (Supplementary Fig. 27). Gelman-Rubin diagnostics plots of the Bayesian stable isotope mixing model.

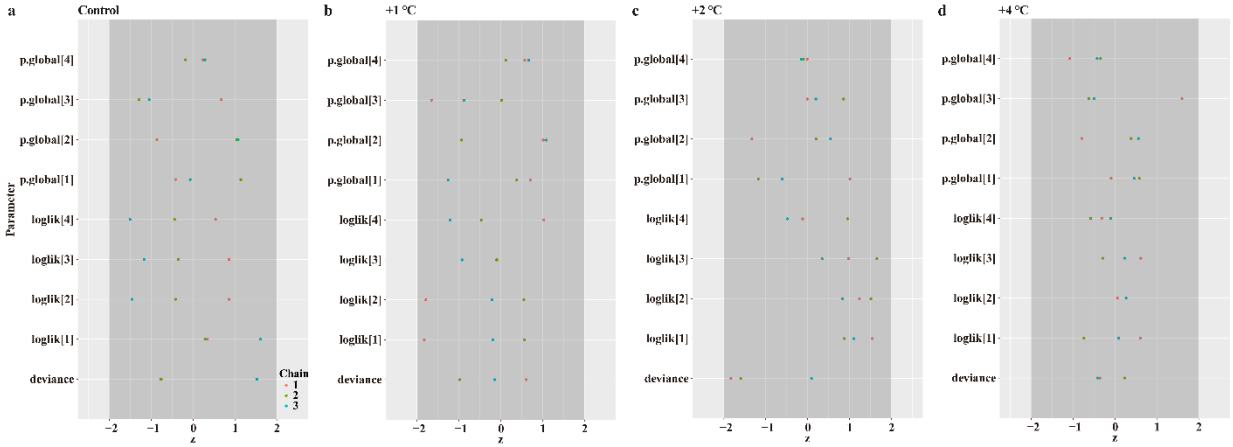

Fig. R12 (Supplementary Fig. 28). Geweke diagnostics plots of the Bayesian stable isotope mixing model.

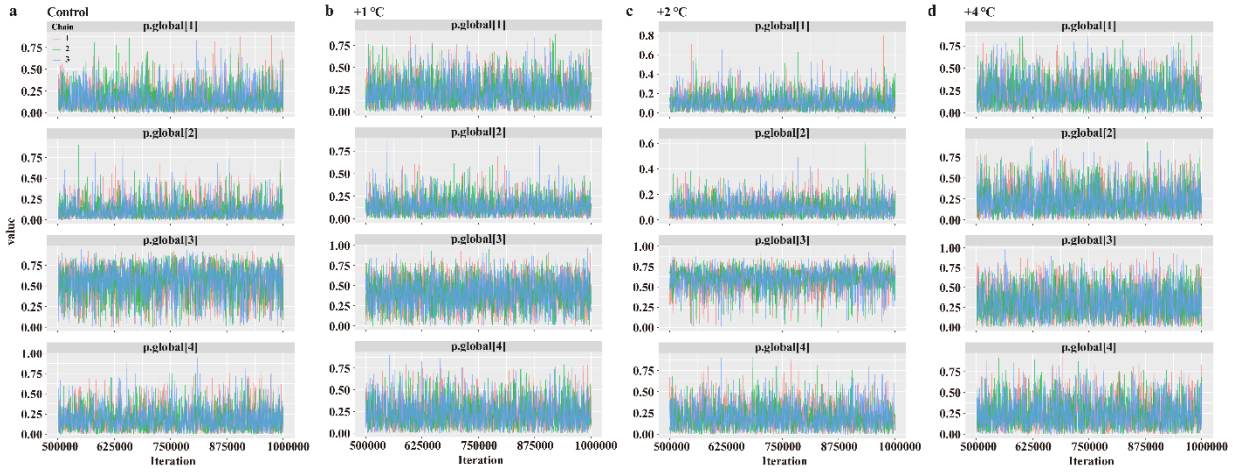

Fig. R13 (Supplementary Fig. 29). Trace plots showing chain mixing of the Bayesian stable isotope mixing model.

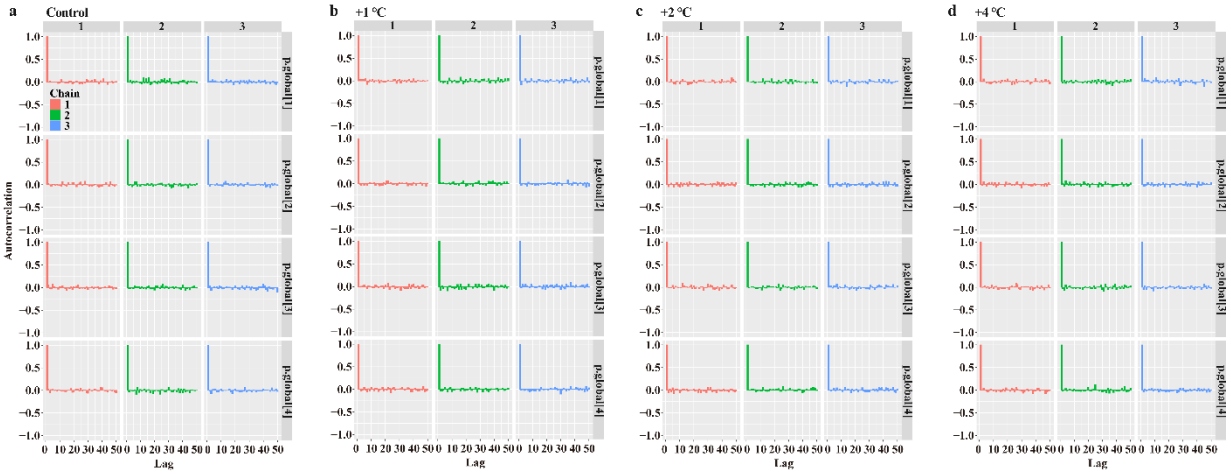

Fig. R14 (Supplementary Fig. 30). Autocorrelation plots of the Bayesian stable isotope mixing model.

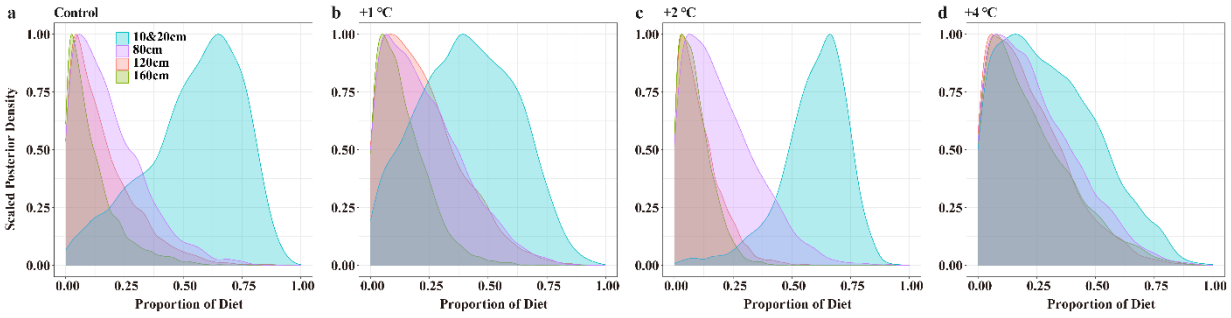

Fig. R15 (Supplementary Fig. 31). Posterior density distributions plots of the Bayesian stable isotope mixing model.

The convergence of all numerical criteria ($\hat{R}$ < 1.01, Geweke z-scores within ±1.96), combined with stable chain behavior and biologically plausible posterior estimates, provides compelling evidence for model reliability despite minor distributional complexities. These comprehensive diagnostics collectively confirm both model convergence and sufficient information to constrain all source contributions. In the Methods, we added the following clarification:

"*Specifically, model convergence was rigorously assessed using Gelman-Rubin diagnostics (all $\hat{R}$ < 1.01; Supplementary Fig. 27), Geweke diagnostics (all |z-score| < 1.96; Supplementary Fig. 28), trace plots (Supplementary Fig. 29), and autocorrelation analysis (Supplementary Fig. 30). Parameter constraints were evaluated through posterior distribution examination (Supplementary Fig. 31).*" (lines 623−627)

[Comment 4] The manuscript reports that the decoupling between Reco and GPP drove a qualitative transition to a strong carbon source, implying a tipping point around 2 °C. However, the temperature gradient in the experiment is uneven, with relatively large gaps between

treatments. While I understand that high-density continuous gradients are logistically challenging, the definition of thresholds in such cases should be more cautious. As these thresholds are used for future projections, the associated uncertainty could be very large and should be explicitly discussed.

Response: Many thanks for this critical and constructive comment. We fully agree that the discrete nature of our experimental warming gradients introduces uncertainty in pinpointing an exact tipping point.

In the updated manucscript, we have removed definitive statements about a tipping point 'around 2 °C' and refer instead to a "*threshold response within approximately 2−4 °C* (lines 374−375)" or a transition from weak to strong carbon sources "*within 2−4 °C warming* (line 418)". These statements better reflect the resolution of our warming treatments.

We have also added a dedicated discussion of this uncertainty, and explicitly caution against the over-interpretation of our threshold for precise future predictions as follows: "*We identify a critical transition in the carbon balance of this permafrost ecosystem within a warming range of ~2−4 °C. Because our warming treatments are discrete, this interval cannot resolve a precise tipping-point temperature. The true threshold likely lies somewhere within this range and may vary among sites. While the result indicates a clear risk of a strong positive carbon-climate feedback emerging within this warming window, its exact numerical value should not be overinterpreted. The central finding is the existence of the transition and its underlying mechanism, that is the warming-induced decoupling of carbon uptake and release driven by old-carbon loss, rather than the exact threshold. Further studies with finer temperature increments are needed to constrain this threshold more precisely.*" (lines 389−397)

Minor issues:

[Comment 5] Line 86, delete ~

Response: Comment accepted.

[Comment 6] Clarify why the results in Fig. 3e and 3f appear contradictory for the 1 °C and 2 °C treatments.

Response: Thanks for this helpful observation. The apparent discrepancy between Fig. 3e and 3f has been resolved by improving our analytical approach. Initially, linear mixed-effects models were used to quantify temperature sensitivity, which may not fully capture the nonlinear responses of *GPP* to temperature. We have now replaced the linear analysis with nonlinear modeling approaches that better represent the biological reality of photosynthetic temperature

responses (**Fig. R16**). With this improved methodology, different methods now show consistent patterns across all warming treatments, including the 1 °C and 2 °C treatments. The revised results demonstrate a consistent response where moderate warming enhances temperature sensitivity, while extreme warming leads to its attenuation.

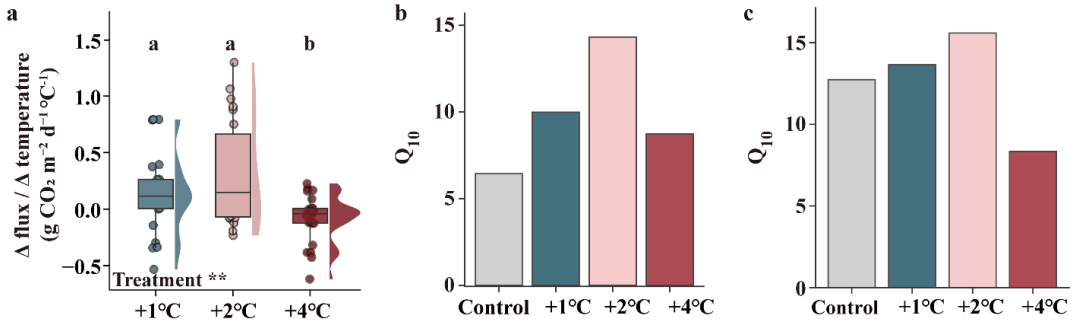

Fig. R16. Temperature sensitivity of gross primary productivity (*GPP*). **a,** Temperature sensitivity represented by change in flux per unit temperature change following multi-level warming. All lowercase letters indicate significant differences between warming treatments. Significant level $*P < 0.05$ and $**P < 0.01$. **b,** Temperature sensitivity (Q_{10}) was quantified using non-linear models (exponential functions). **c,** Temperature sensitivity (Q_{10}) was quantified by isolating the net warming effect with non-linear models (exponential functions) that incorporated random effects and controlled for soil moisture covariation.

We have updated the calculation method of temperature sensitivity in the Methods section accordingly as follows: "*The temperature sensitivity of CO_2 fluxes was quantified using three complementary approaches. First, we calculated the direct sensitivity metric, defined as the change in flux (Δflux) per 1 °C increase relative to the control, to represent the net warming effect under different treatments. Second, we assessed the temporal temperature sensitivity (Q_{10}) using exponential functions by analyzing the time series relationships between fluxes and temperature across warming levels, as follows (Eq. 2, Eq. 3*[62]*):*

$$R = ae^{bt} \quad (2)$$

$$Q_{10} = e^{10b} \quad (3)$$

where R represents CO_2 flux, a denotes the flux at 0 °C, b denotes the model fitting parameter, t denotes the soil temperature at the 10 cm layer, and Q_{10} indicates the temperature sensitivity, that is, the multiplicative increase in the CO_2 flux for every 10 °C increase in soil temperature. Third, because humidity fluctuations can influence temperature responses, we conducted partial correlation analyses using nonlinear models (exponential functions or generalized additive mixed models (GAMMs)) controlling for humidity to isolate and quantify the temperature-driven

component of flux variation. Using the fitted GAMMs, we derived a dynamic Q_{10} by calculating the flux ratio (Flux(T+10) / Flux(T)) for each 1 °C step within the observed temperature range. We then presented the average Q_{10} for each treatment over this range." (lines 632−646) We also made corresponding revisions in the Results and Discussion sections (lines 137−138 and lines 231−233). We believe this revision strengthens our study by employing more biologically realistic temperature response functions.

[Comment 7] The unit of GPP is incorrect; as a rate, it should include a time component.

Response: We have corrected the presentation of the *GPP* units throughout the manuscript to address this point. Thank you for this correction.

[Comment 8] In Fig. 5, under Phase III, clarify why the absolute value of ΔGPP is greater than the absolute value of ΔER.

Response: Insightful question. First, we clarify that the values of Δ*GPP* and ΔR_{eco} describe the net change within each specific phase, rather than cumulative changes since the start of the experiment.

In Phase III, *GPP* experiences an abrupt decline, while R_{eco} shows only a modest increase. When calculated solely over this phase, the magnitude of *GPP* suppression (|Δ*GPP*|) is therefore greater than the magnitude of the R_{eco} increase (|ΔR_{eco}|). New Supplementary Fig. 19 (**Fig. R17**) presents the stepwise changes across warming levels (Control → +1 °C, +1 °C → +2 °C, +2 °C → +4 °C), confirming that the +2 °C to +4 °C transition (Phase III) indeed yields a large negative Δ*GPP* (147.19 g CO_2 $m^{-2}yr^{-1}$, 29% decline) and a small positive ΔR_{eco} (17.14 g CO_2 m^{-2} yr^{-1}, 2.12% increase), consistent with our interpretation.

Mechanistically, this asymmetry arises because extreme warming directly suppresses *GPP* through physiological stress, likely exceeding the thermal optimum of dominant alpine species, leading to a sharp nonlinear reduction in photosynthetic capacity. In contrast, the increase in R_{eco} is moderated by two counteracting processes: (i) enhanced decomposition of thaw-released old carbon, and (ii) diminishing labile substrates and reduced contemporary carbon input under declining *GPP*. As a result, R_{eco} increases only slightly, while *GPP* declines strongly, yielding |Δ*GPP*| > |ΔR_{eco}| in Phase III. This imbalance drives the substantial net CO_2 efflux observed. We have incorparated these points into Discussion (lines 295−299).

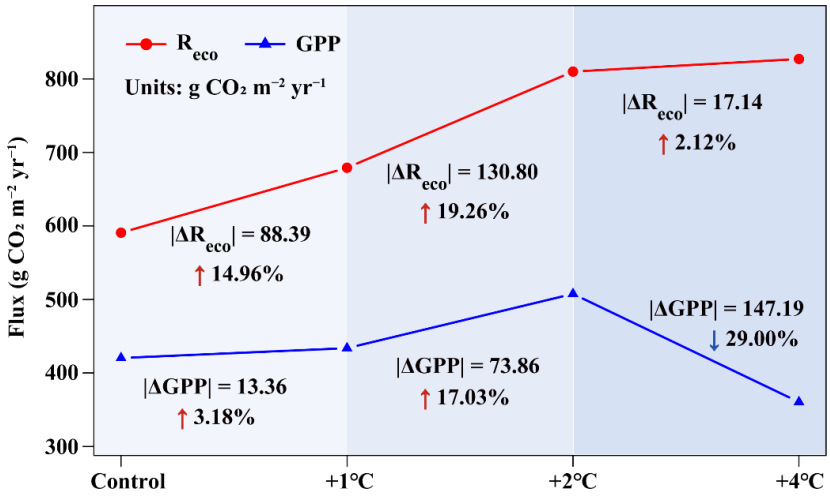

Fig. R17. Changes of gross primary production (*GPP*) and ecosystem respiration (R_{eco}) across warming gradients during the experimental periods. Red circles and lines represent R_{eco}, while blue triangles and lines represent *GPP*. Units for all values are in g CO_2 m^{-2} yr^{-1}. The red upward arrow indicates an increase in flux, and the blue downward arrow indicates a decrease.

We have added the following text to the caption of Fig. 5 as follows: "*Note that ΔGPP and ΔR_{eco} represent the changes within each discrete warming phase, not the cumulative change from the initial state.*" (lines 316–317) and "*The change from +2 °C to +4 °C (corresponding to Phase III) exhibits a large negative ΔGPP ($|\Delta GPP|$ = 147.19 g CO_2 m^{-2} yr^{-1}, indicating a 29% decline) and a small positive ΔR_{eco} ($|\Delta R_{eco}|$ = 17.14 g CO_2 m^{-2} yr^{-1}, indicating a 2.12% increase; Supplementary Fig. 19), consistent with our schematic (Fig. 5a).*" (lines 327–329)

[Comment 9] Line 134(also other places), change "oC" to "°C" in the unit g CO_2 m^{-2} d^{-1} $°C^{-1}$.

Response: We have corrected the temperature unit from "oC" to the proper symbol "°C" throughout the entire manuscript, including in the unit "g CO_2 m^{-2} d^{-1} $°C^{-1}$" on Line 134 and all other places. Thank you for the suggestion.

[Comment 10] Maintain consistent use of significance symbols across figures (lowercase letters vs. asterisks) and explain the symbol scheme in each figure legend.

Response: We have standardized the significance notation across all figures by using lowercase letters. Each figure legend explicitly explains the notation scheme: "*All lowercase letters indicate significant differences between warming treatments.*" (lines 112–113, lines 142–143, and lines 148–149)

[Comment 11] Double check the unit of GPP and ER in the manuscript are g CO2 not g C, usually GPP in QTP is about 200 to 900 g C m^{-2} yr^{-1}, if it is g CO2, it should be much larger.

Response: Thank you for this nice comment. We have double-checked our data processing and unit reporting, and we confirm that the units for *GPP* and R_{eco} are correctly presented as g CO_2 m^{-2} yr^{-1}.

We agree that our reported *GPP* values are lower than the typical range cited by the reviewer (200−900 g C m^{-2} yr^{-1}, which is equivalent to ~733−3300 g CO_2 m^{-2} yr^{-1}). However, these differences are consistent with the environmental constraints at our study site. Specifically, the site is located at ~4790 m elevation, with a short growing season (<100 days) and low mean summer temperatures (4.97 ± 0.44 °C[4]), all of which strongly reduce plant productivity and microbial activities. Therefore, the relatively low *GPP* values reported here reflect the harsh climatic conditions rather than unit inconsistencies.

REFERENCES

1. Shao, P. et al. Tradeoffs among microbial life history strategies influence the fate of microbial residues in subtropical forest soils. *Soil Biol. Biochem.* **153**, 108112 (2021).
2. Smith, C. et al. The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity Supplementary Material. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (eds. Masson-Delmotte, V. et al.) (Cambridge University Press, 2021).
3. Wang, G. et al. Enhanced response of soil respiration to experimental warming upon thermokarst formation. *Nat. Geosci.* **17**, 532–538 (2024).
4. Yang, K. et al. China meteorological forcing dataset (1979–2018). National Tibetan Plateau / Third Pole Environment Data Center (2019).

RESPONSE TO REVIEWERS' COMMENTS

We sincerely appreciate the two reviewers' great efforts in reviewing our manuscript and, especially, the constructive comments and insightful suggestions for further improving the manuscript's quality. Following is our point-by-point response to the reviewers, and the manuscript has been updated accordingly (with changes marked in *red*). For clarity and easier navigation, referee comments are presented in blue, while our responses are in black.

Reviewer #1

The authors have done a great job addressing all the reviewer's comments and were very considerate, detailed and thoughtful in their responses. It seems that the manuscript has clearly benefited from the reviewer's comments as it is much easier to understand, more clearly written and focused and more direct evidence. I just have several minor comments:

Response: Thank you very much for your supportive evaluation and constructive comments that have substantially improved our manuscript. We have carefully addressed all concerns raised and revised the paper accordingly. Below, we provide point-by-point responses to all comments, and we hope our revision is to your satisfaction.

[Comment 1] L60-To be consistent, cm a-1?

Response: Comment accepted.

[Comment 2] L457-The target warming magnitude is +1 °C, +2 °C, and +4 °C. So what is the actual temperature increase, and should be indicated and clarified.

Response: Many thanks for this insightful comment. The experimental treatments were designed to simulate specific climate warming scenarios: target increases of +1 °C, +2 °C, and +4 °C relative to the ambient control. To achieve these precise scenarios, we employed a precisely controlled infrared warming system. This system features an automatic feedback mechanism that continuously monitors the air temperature and adjusts the heating output in real-time, ensuring that the target warming magnitudes were accurately achieved at the control point. This precise control allowed us to reproduce the designated climate warming scenarios with high accuracy, essentially reaching the target warming magnitudes of +1 °C, +2 °C, and +4 °C. Under the driving force of these simulated climate scenarios, the soil temperatures across the entire profile

increased significantly (all $P < 0.001$). The magnitude of soil warming naturally attenuated with depth due to thermal conductivity but remained statistically significant at all layers (**Table R1**). For instance, at 10 cm depth, the mean soil temperature increases relative to the control were approximately +1.3 °C, +1.6 °C, and +3.0 °C for the +1 °C, +2 °C, and +4 °C air warming scenarios, respectively. At 160 cm depth, the corresponding increases were +0.3 °C, +0.4 °C, and +0.9 °C. We have now added Supplementary Table 2 detailing the actual soil temperature increases (**Table R1**) and incorporated a clarification in the Methods section accordingly as follows: "*We established a multi-level warming experiment using precisely controlled infrared heaters with four treatments: ambient temperature (Control) and target increases of +1 °C, +2 °C, and +4 °C above ambient to simulate specific climate warming scenarios. This warming system features an automatic feedback mechanism that continuously monitors the air temperature and adjusts the heating output in real-time, ensuring that the target warming magnitudes were accurately achieved at the control point. This precise control allowed us to reproduce the designated climate warming scenarios with high accuracy, essentially reaching the target warming magnitudes of +1 °C, +2 °C, and +4 °C. The treatments were replicated four times in a randomized block design across sixteen plots (each plot covered an effective observation area of approximately 7 m^2).*" (lines 461−469) and "*Under the driving force of these simulated climate scenarios, the soil temperatures across the entire profile increased significantly (all $P < 0.001$; Supplementary Table 2). The magnitude of soil warming naturally attenuated with depth due to thermal conductivity but remained statistically significant at all layers. For instance, at 10 cm depth, the mean soil temperature increases relative to the control were approximately +1.3 °C, +1.6 °C, and +3.0 °C for the +1 °C, +2 °C, and +4 °C air warming scenarios, respectively. At 160 cm depth, the corresponding increases were +0.3 °C, +0.4 °C, and +0.9 °C.*" (lines 483−488).

Depth	Control	+1 °C	+2 °C	+4 °C	*P*
10 cm	0.94 ± 0.02 d	2.25 ± 0.02 c	2.55 ± 0.02 b	3.89 ± 0.02 a	< **0.001**
20 cm	1.09 ± 0.01 d	1.90 ± 0.01 c	2.23 ± 0.01 b	3.40 ± 0.02 a	< **0.001**
80 cm	1.14 ± 0.01 d	1.65 ± 0.01 c	1.77 ± 0.01 b	2.62 ± 0.01 a	< **0.001**
120 cm	1.08 ± 0.01 d	1.52 ± 0.01 c	1.60 ± 0.01 b	2.29 ± 0.01 a	< **0.001**
160 cm	1.01 ± 0.01 d	1.34 ± 0.01 c	1.42 ± 0.01 b	1.91 ± 0.01 a	< **0.001**

Table R1 (Supplementary Table 2). The average temperature across the soil profile under the gradient warming treatment (mean ± standard error). Unit: °C. Values within the same row followed by different letters indicate significant differences between warming treatments.

[Comment 3] It is recommended to strengthen the link between the Discussion and the two scientific hypotheses presented in the Introduction. For example, are the hypotheses supported? Why or why not?

Response: Following your suggestion, we have added explicit statements evaluating these hypotheses in the Discussion to clarify these points accordingly as follows: "*Our multi-gradient warming experiment reveals nonlinear temperature responses of key ecosystem carbon exchange processes (Fig. 5b), providing direct evidence supporting hypothesis (i) that incremental warming induces asymmetric peaks and eventual decoupling of ecosystem respiration (R_{eco}) and photosynthetic carbon fixation (GPP).*" (lines 228−231) and "*Crucially, the "metabolic valve" mechanism amplifies under 4 °C warming (enhanced to 76% contribution), driving sustained high R_{eco} levels, which strongly supports our second hypothesis. It confirms that the nonlinear surge in R_{eco} is disproportionately fueled by old carbon decomposition, which critically underpins the shift of the ecosystem from a carbon sink to a source.*" (lines 386−390).

[Comment 4] Supplementary Fig. 27- all $\hat{R}$ is equal to 1? These indicators require more explanation for understanding. Same for Supplementary Fig. 28

Response: Thank you for this comment. Gelman-Rubin diagnostic ($\hat{R}$) compares the variance between multiple MCMC chains to the variance within each chain. An $\hat{R}$ value close to 1 (typically < 1.05) indicates that the chains have converged to the same target distribution. Our results clearly show the actual point estimates $\hat{R}$ for each parameter were all well below the threshold of 1.05, indicating excellent convergence. Geweke diagnostic (z-score) compares the mean of the early part of an MCMC chain to the mean of the later part. A |z-score| < 1.96 suggests no significant difference between these segments, indicating the chain has reached a stationary state. All our |z-scores| were < 1.96.

We have added details to the Methods section accordingly as follows: "*Specifically, model convergence was rigorously assessed using: (1) Gelman-Rubin diagnostics (all $\hat{R}$ < 1.01; Supplementary Fig. 27), where $\hat{R}$ compares between-chain to within-chain variance, with values near 1 (typically < 1.05) indicating convergence; (2) Geweke diagnostics (all |z-score| < 1.96; Supplementary Fig. 28), where the z-score tests stationarity by comparing the early and late segments of the chain; (3) visual inspection of trace plots (Supplementary Fig. 29); and (4) autocorrelation analysis (Supplementary Fig. 30). Parameter constraints were evaluated through posterior distribution examination (Supplementary Fig. 31).*" (lines 639−645).

Reviewer #2

From Editor: Please know Reviewer #2 finds that the current version has addressed all concerns and that the manuscript has been thoroughly improved.

Response: We appreciate your positive evaluation and kind encouragements on our work.